VARIATION OF THE FREQUENCY OF FATAL QUARRELS WITH MAGNITUDE

LEWIS F. RICHARDSON
Hillside House, Kilmun, Argyll, Britain

A record of wars during the interval A.D. 1820 to 1945 has been collected from the whole world, and has been classified according to the number of war-dead. The smaller incidents have been the more frequent, according to a fairly regular graph which can be extended to quarrels that caused a single death.

INTRODUCTION

ANYONE WHO TRIES to make a list of "all the wars" (e.g. Wright [18] p. 636) encounters the difficulty that there are so many small incidents, that some rule has to be made to exclude them.

From the psychological point of view a war, a riot, and a murder, though differing in many important aspects, social, legal, and ethical, have at least this in common that they are all manifestations of the instinct of aggressiveness. This Freudian thesis has been developed by Glover [5], by Durbin & Bowlby [3], and by Harding [6]. There is thus a psychological justification for looking to see whether there is any statistical connection between war, riot, and murder. I dealt with this problem in 1941 by forming the inclusive concept of a "fatal quarrel," and then classifying the fatal quarrels according to the number of quarrel-dead [12]. By a *fatal quarrel* is meant any quarrel which caused death to humans. The term thus includes murders, banditries, mutinies, insurrections, and wars small and large; but it excludes accidents, and calamities such as earthquakes and tornadoes. Deaths by famine and disease are included if they were immediate results of the quarrel, but not otherwise. In puzzling cases the legal criterion of "malice aforethought" was taken as a guide.

The record of the number killed in a particular war is often uncertain by a factor of two. The meaningful part of the record can be separated from its uncertainty by taking the logarithm to the base ten and rounding the logarithm off to a whole number, or to the first decimal, according to the quality of the information. For simplicity I have lumped together the deaths on the opposing sides of the quarrel. The *magnitude* of a fatal quarrel is defined to be the logarithm to the base ten of the number of people who died because of that quarrel. The magnitude will be denoted by μ. The range of magnitude extends from 0 for a murder involving only one death, to 7.4 for the Second World War. Other well-known wars had magnitudes as follows: 1899–1902 British versus Boers 4.4; 1939–40 Russians versus Finns 4.83; 1861–65 North American Civil War 5.8. The magnitude of a war is usually known to within ± 0.2; so that a classification by unit ranges of magnitude is meaningful. These ranges have been marked off at 7.5, 6.5, 5.5, 4.5, 3.5, 2.5, 1.5 and perhaps at 0.5 and -0.5. Every fatal quarrel must lie inside one of these cells; it cannot lie on a boundary, because the antilogs of the boundary μ are not integers.

This abstract framework becomes of interest when the facts are sorted into it. As far as I have been able to ascertain, the historians have never sorted their facts by the scale of magnitude. The chief obstacle to any scientific study of wars-and-how-to-avoid-them, is each nation's habit of blaming other nations. National prejudice can however be avoided by taking the whole world as the field of study; and by taking a time-interval longer than personal memory. For this reason I have made a search for the records of fatal quarrels in the whole world since the beginning of A.D. 1820. The task has been long; and the results are here presented in brief summary. For magnitudes greater than 2.5, the facts were mostly obtained from works on history. For magnitudes less than 0.5 they were taken from criminal statistics. For magnitudes between 0.5 and 2.5 the information is scrappy and unorganized; what there is of it suggests that such small fatal quarrels were too numerous and too insignificant to be systematically recorded as history, and yet too large and too political to be recorded as crime.

It is important to know whether the facts have all been gathered. The best evidence is provided by the progress of search. I began in the year 1940. At first my collection grew rapidly, then slowly; then various revisions caused its totals to oscillate slightly. The collection appeared to be sufficiently complete to warrant the publication of a summary, which appeared in *Nature* of 15th November, 1941. Afterwards the publication of Quincy Wright's list of wars ([18] Appendix XX) provided a stimulus to further enquiry, which involved the consultation of some seventy history-books. In the following table three stages of my records are compared. All refer to fatal quarrels which ended from A.D. 1820 to 1929 inclusive.

Date and place of record	Ends of range of magnitude			
	$7\pm\frac{1}{2}$	$6\pm\frac{1}{2}$	$5\pm\frac{1}{2}$	$4\pm\frac{1}{2}$
	Numbers of fatal quarrels recorded			
1941. *Nature* of Nov. 15	1	3	16	62
1944. J.R.S.S. 107, 248	—	—	—	63
1947. Here.	1	3	20	60

It is seen that there has been a sort of convergence, which makes the present numbers worthy to be discussed. That is for $\mu > 3.5$. For smaller incidents in the range $2.5 < \mu < 3.5$ my collection has continued to grow; the sign $\geqq$ is therefore prefixed to its totals.

The world-total for the murders was obtained from the murder-rates of 17 countries. They were weighted by populations. The weighted mean rate for the world was found to be 32 murders per million per year. This is not as definite as the statistics of wars, because the criminal statistics of China, Africa, and South America (except Chile) were not available. But even if that estimate were as much as three times wrong, which is incredible, it still would give significant results in Figures 0 and 1, because they relate to such enormous ranges. The evidence which I had concerning time-changes in the murder-rate was scrappy and conflicting; so that provisionally I took the rate to be constant. The mean population of the world over the 126 years from A.D. 1820 to 1945 was computed from a publication by Carr-Saunders [1A] and was found to be 1.49×10^9. So the world-total of murders was computed as $1.49\times10^9\times32\times10^{-6}\times126=6.0\times10^6$.

A revision based on the murder-rates of 21 countries, gave a mean rate of 37 instead of 32, and a world-total of 7 instead of 6 million. The following diagrams and tables were based on the earlier estimate and have not been altered.

A CONSPECTUS WHICH SUITS THE WARS

The number of fatal quarrels has been counted, or estimated, in unit ranges of magnitude, with the following results for the world as a whole. The date of a quarrel is taken to be that of the termination of hostilities. The interval is A.D. 1820 to 1945 inclusive.

Ends of range of magnitude	$7\pm\frac{1}{2}$	$6\pm\frac{1}{2}$	$5\pm\frac{1}{2}$	$4\pm\frac{1}{2}$	$3\pm\frac{1}{2}\cdots$ Murders
Observed number of fatal quarrels	2	5	24	63	$\geqq 188 \cdots 6\times10^6$

It is desirable to show all these facts together on a single diagram. If the number of fatal quarrels per unit range of magnitude were taken as the ordinate, then either the wars would show, or the murders, but not both. When however the logarithm of the number of fatal quarrels per unit range of magnitude is taken as the ordinate, then all parts show equally well. Each murder presumably involved a small number of deaths, such as 1, 2, 3, or 4, so that the corresponding logarithms were 0.000, 0.301, 0.477, or 0.602. How to group them was not obvious. As a first expedient I extended to the murders the system of classification which had been found to suit the wars, namely by unit ranges of magnitude cut at 7.5, 6.5, 5.5, 4.5, 3.5, and so on by extension to 0.5 and -0.5. Murders involving 4 or more deaths were regarded as negligibly rare. The ordinate f is defined to be

$$f = \log_{10} \left\{ \frac{\text{number of fatal quarrels}}{\text{corresponding range of magnitude}} \right\}.$$

This rough and ready scheme gives the conspectus shown in Figure 0. It is the type of diagram with which the author began, and it has now been asked for by a referee. A dotted curve has been drawn across the gap to suggest that there is statistical continuity between the wars and the murders.

THE FAULTS OF FIGURE 0 AND HOW TO AMEND THEM

In order to match the wars, the murders have been enclosed in a unit range of magnitude, as shown at A in Figure 0. But the half of this range, from $\mu = -0.5$ to μ just less than zero, is necessarily empty. Would it be more reasonable to regard the murders as enclosed in a half-unit range? If so the line representing them in Figure 0 would be halved in length. Also the number of murders per unit of magnitude would be doubled. So the representative line would be raised by $\log_{10} 2$ into position marked B. Worse ambiguities occur if we wish to show separately the number of murders each causing one death. Let there be M of them. To what range of magnitude ought they to be attributed? The range must contain the point $\mu = 0$, and must not contain the point $\mu = \log_{10} 2$, but otherwise is unspecified. Yet M has to be divided by this indefinite range in order that the reckoning may be "per unit range of magnitude". Figure 0, though suitable for the wars, is too vague for quarrels that caused only one death.

It may be thought that the root of the trouble is the reckoning "per unit range of" anything, and that if ranges were abandoned then murders and wars could be compared by plotting the logarithm of their numbers against their magnitudes. Such a diagram would consist of

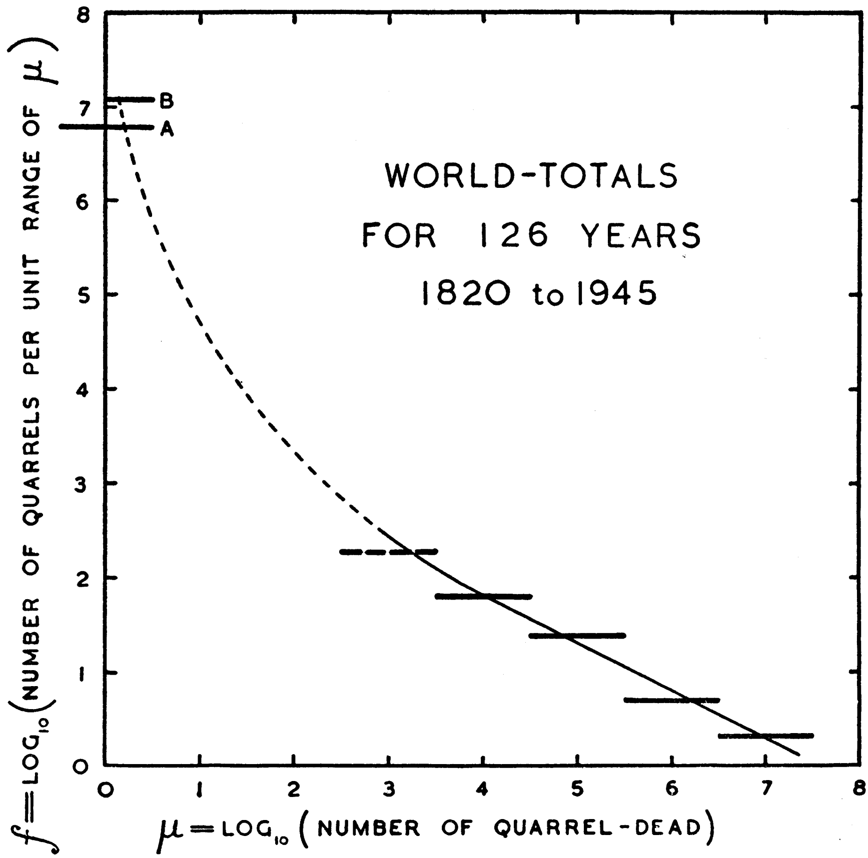

FIGURE 0

points, not segments. There are two objections to this proposal: (i) the magnitudes of wars are imperfectly known and some of them are rounded off to whole numbers; so that there would be many artificially coincident points, (ii) if the magnitudes were better known they would presumably be scattered, so that a point-diagram of wars would present great irregularities, of no interest. For when n exceeds 1000 there is never likely to be any statistically significant distinction between wars involving n and $n+1$ deaths. In brief we must have points for the murders and averages over ranges for the wars, and yet these diverse representations must somehow be compared. The solving idea, which melted all the obstacles, was that the appropriate measure of any range of quarrel-dead is simply the number of integers in that range, and not the corresponding range of magnitude, nor any other artificial conception.

A DIAGRAM WHICH SUITS THE MURDERS

The logically simple idea is that there were, say $q(n)$ quarrels each involving n deaths. (1)

For example in England and Wales during 1935 and 1936 there were known to the police 186 cases of murder of 213 persons aged over one year. In 74 cases (involving 91 persons) the murderer or suspect committed suicide. In 105 cases (115 victims) 111 persons were arrested. In the remaining 7 cases, involving 7 victims, no arrest was made. Of the 111 persons arrested 16 were executed (Whitaker's Almanacks 1938 and 1939, p. 652). These statements can be interpreted, with only slight uncertainty, by the following distribution.

			totals
$n=1$	2	3	
$q(n)=92$	71	23	186
$nq(n)=92$	142	69	303

The number of deaths, n, connected with the quarrel is here taken to include, along with those of the victims, also those of the murderer or suspect, whether by suicide or execution. It seems permissible, for illustration until better information can be found, to regard the 6 million murders in the world from 1820 to 1945 as the number of fatal quarrels for $n=1$, 2 and 3, and to subdivide this total among its three parts in the same ratios as the total for England and Wales is subdivided. Thus one finds for the world, the following enlarged values of $q(n)$. To take one country as a sample of the world is very unsatisfactory. But it should be noted that the uncertainties of these $q(n)$ are insignificant in comparison with the vast range of Figure 1 where $q(n)$ varies in a ratio exceeding 10^{13}.

			totals
$n=1$	2	3	
$q(n)=3.0$	2.3	0.7	6.0 millions
$nq(n)=3.0$	4.6	2.1	9.7 millions

These give three points in the top left corner of Fig. 1, with the coordinates

$\log_{10} n=0$	0.301	0.477
$\log_{10} q(n)=6.48$	6.36	5.85

A GENERAL AND WELL DEFINED CONSPECTUS

This will now be obtained by extending to the wars the method introduced above for the murders.

In the region of wars, say for $n > 1000$, the function $q(n)$ is usually zero and occasionally unity; so that a diagram of it would be bristly and unreadable. Therefore for the wars we should take an average of $q(n)$, say $\bar{q}(n)$. As a preliminary to the definition of $\bar{q}(n)$, it must be understood that the fatal quarrels are arranged in order of n. The definition of $\bar{q}(n)$ is then

$$\bar{q}(n) = \frac{\text{number of fatal quarrels in a range of } n}{\text{number of integers in that range of } n} \cdot \tag{2}$$

Briefly, and somewhat inaccurately, $\bar{q}(n)$ may be called the "average frequency of quarrels per unit range of quarrel-dead." As in all frequency-diagrams so here the choice of the group-range has to be a compromise: if the range is too small, the diagram is spiky; if the range is too large, essential features are flattened out. The question of the best range must be decided by trial.

The diagram of q and $\bar{q}$ as functions of n is the logically simple conspectus of the frequency of wars and murders. This diagram cannot be drawn, because no sheet of paper is large enough to show the facts at both ends of the range. The difficulty is overcome by plotting the logarithms of q, $\bar{q}$ and n. The magnitude μ has already been defined by

$$\mu = \log_{10} n. \tag{3}$$

Another oft-recurring symbol, ϕ, must now be defined.

Let $\log_{10} \bar{q}(n) = \phi(\mu)$, say, for the wars: while for the murders

$$\phi(\mu) = \log_{10} q(n) \text{ without averaging.} \tag{4}$$

The last column of the following table shows $\phi(\mu)$, apart from some corrections for grouping, which will be explained on p. 542. The first

Range of magnitude μ	Number of integers in range of war-dead	Observed number of fatal quarrels Years 1820–1945	$\log_{10}$ (number of quarrels per unit range of war-dead)
2.5 to 3.5	2,846	$\geq$188	≥ -1.180
3.5 to 4.5	28,460	63	-2.655
4.5 to 5.5	284,605	24	-4.074
5.5 to 6.5	2,846,050	5	-5.755
6.5 to 7.5	28,460,499	2	-7.153

and last columns of this table provide the abscissa and ordinate of the accompanying diagram (Figure 1), as far as the wars are concerned. It is evident that the wars can be satisfactorily fitted by a straight sloping line. On account of grouping, the wars are represented, not by points as the murders are, but by horizontal segments. It will be shown later (p. 542) that the $\phi(\mu)$ graph should pass slightly above the midpoints of these segments. That is to say the sloping straight line was actually obtained by a second approximation.

The relation between the ordinates of Figures 0 and 1

The abscissa μ is the same in both diagrams and the ordinates are respectively f and ϕ, defined thus

$$f = \log_{10}\left\{\frac{q(n_1) + q(n_1+1) + q(n_1+2) + \cdots + q(n_1+s)}{\log_{10} n_b - \log_{10} n_a}\right\}$$

where $n_1 - 1 < n_a < n_1$, and $n_1 + s < n_b < n_1 + s + 1$. These extra bits at the ends of the range have to be put in to prevent f becoming infinite at $s = 0$.

$$\phi = \log_{10}\left\{\frac{q(n_1) + q(n_1+1) + q(n_1+2) + \cdots + q(n_1+s)}{s+1}\right\}$$

which is definite at $s = 0$. By subtraction

$$f - \phi = \log_{10}\left\{\frac{s+1}{\log_{10} n_b - \log_{10} n_a}\right\}.$$

This expression involves an ambiguity, because n_b is loose in a range of unity, and so is n_a. The practical question however is whether the ambiguity is greater than the uncertainty of the observations. It is necessary to distinguish different cases.

Case (i). The range here used in treating the wars is $\mu_0 - \frac{1}{2} \leqq \mu \leqq \mu_0 + \frac{1}{2}$ where $\mu = \log_{10} n$ and where $\mu_0 \geqq 3$. In these circumstances arithmetic shows that

$$f - \phi = \mu_0 + 0.454, \tag{5}$$

reliable to the third decimal. Seeing that the μ of a particular war is often uncertain by ± 0.1, equation (5) is admirably accurate for the purpose.

Case (ii). We may conceivably wish to consider a shorter range $\mu_0 - \epsilon \leqq \mu \leqq \mu_0 + \epsilon$ where μ_0 is still in the region of the wars. If ϵ is small, but not too small, there is a quasi-limit such that

$$f - \phi = \mu_0 + 0.362. \tag{5A}$$

This was obtained by approximating to the difference-ratio by a derivative, and was confirmed by arithmetic.

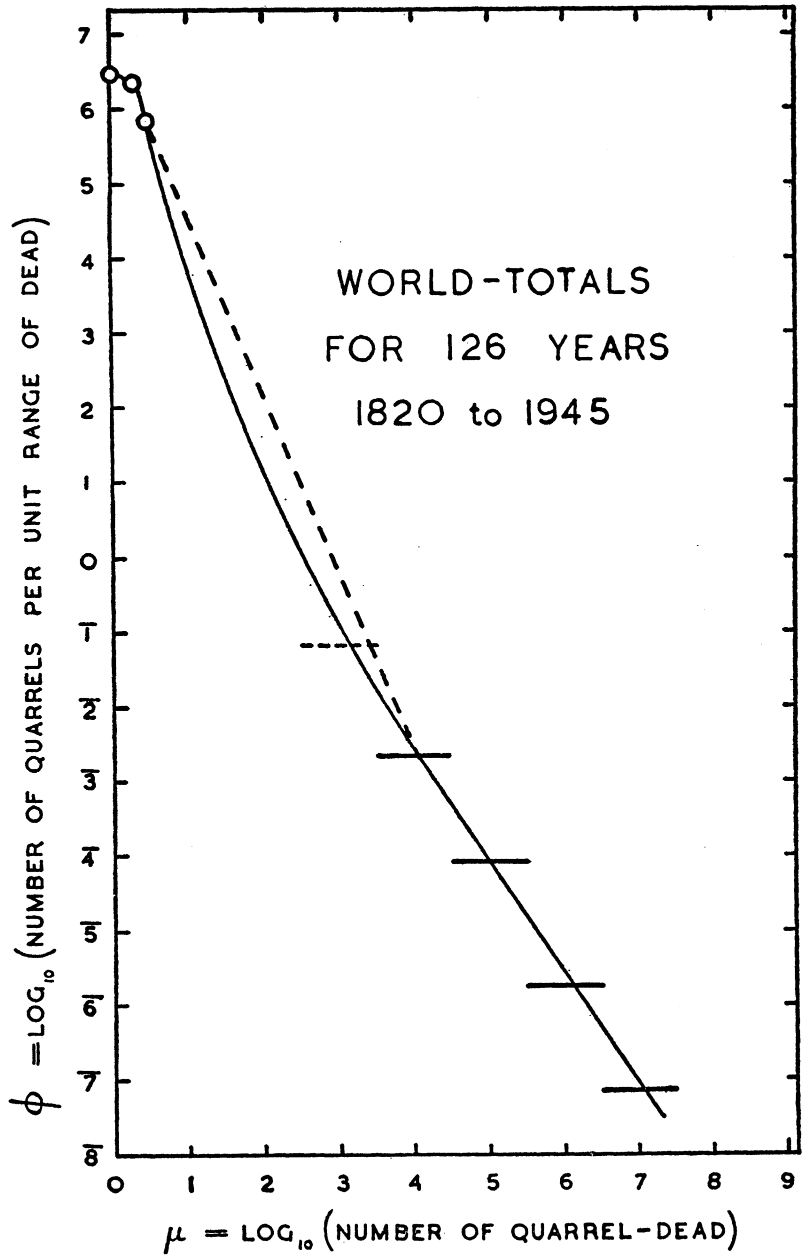

FIGURE 1. THE WHOLE RANGE OF FATAL QUARRELS

THE HORIZONTAL SCALE EXTENDS FROM A SINGLE PERSON ON THE LEFT TO THE POPULATION OF THE WORLD ON THE RIGHT. THE THREE DOTS AT THE LEFT HAND TOP REPRESENT THE MURDERS. THE SHORT HORIZONTAL LINES REPRESENT THE WARS GROUPED BY MAGNITUDES. ONE OF THESE LINES IS DOTTED, BECAUSE IT SHOULD BE RAISED TO AN UNKNOWN HEIGHT.

Case (iii). If there are only a few integers n in the range then the ambiguity becomes troublesome. This is the situation as to the murders. The ambiguity of $f-\phi$ is more than that of some criminal statistics. The fault lies in f, not in ϕ. When $n_1=1$ and $s=0$ it is hardly possible to locate n_a and n_b by any reasonable convention, as already explained on page 526. To revert to Cases (i) and (ii), either of them give the same relation between the slopes of the f- and ϕ-diagrams. As these slopes are in fact negative, it is convenient to write the relation thus

$$-\, d\phi/d\mu_0 = -\, df/d\mu_0 + 1 \tag{5B}$$

which shows that the ϕ-diagram is the steeper.

The slope in the region of the wars

This was found, by the principle of Maximum Likelihood, [4 & 10] to be

$$d\phi/d\mu_0 = -\, 1.50. \tag{5C}$$

Interpolation across the gap between the wars and the murders

The dotted horizontal segment for the small wars represents only a lower bound, and otherwise should be ignored. Between the definite end-points of the gap, the mean slope is

$$d\phi/d\mu = -\, 2.38. \tag{6}$$

This is shown by a straight line of dashes.

As to the choice of a curve, there are the general clues of simplicity and of continuous slope. For simplicity a circular arc was chosen; and, for continuous slope, the arc was made tangent to the straight line which fits the wars. It was then found, as an independent confirmation that the slope of the circular arc joins on continuously to that for the murders. Some collateral evidence is provided by banditry.

A SEARCH FOR THE FACTS ABOUT QUARRELS THAT CAUSED FROM 4 TO 315 DEATHS $(0.5<\mu<2.5)$

That there were many small fatal quarrels is shown by a statement made by Quincy Wright [18] in the introduction to his long list of "Wars of Modern Civilization." He remarked that "A list of all revolutions, insurrections, interventions, punitive expeditions, pacifications, and explorations involving the use of armed force would probably be more than ten times as long as the present list, . . . " Wright however did not classify fatal quarrels according to the number of quarrel-dead. Similarly whilst reading histories and police-reports and listening to radio-news I have noticed allusions to very many incidents less important than wars, but more important than murders. References [2], [7], [11], [15], contain some remarkable examples. However I see no hope of obtaining world-totals of the numbers of such incidents during the 126 years to which Figure 1 relates, and classified moreover by magnitudes in the range $0.5<\mu<2.5$. Any factual test of the interpolated portion of Figure 1 will have to depend on smaller samples. The record of banditry in Manchoukuo is comparable on certain assumptions, and that of gangsters in Chicago on further assumptions. These ordered collections of facts, being rare specimens, are correspondingly valuable.

BANDITRY IN MANCHOUKUO DURING THE YEAR 1935

This is of especial interest, because the facts are given in a form which throws light on aggregation for aggression. The following is quoted from the *Japan and Manchoukuo Year Book*, 1938, pp. 692–95.

"At the time of the founding of Manchoukuo in March of 1932 the total number of bandits exceeded 100,000. By September of the same year the number had increased to 210,000 due principally to the subversive activities of Chang Hsueh-Liang's remnant troops who were thrown out of employment following the downfall of the young marshal. Since then, however the number of such bandits has been on the decrease as a result of their suppression by Manchoukuo and Japanese forces.

"Compared with the condition obtaining in 1932 two factors loom in prominence with regard to the bandit situation. Firstly may be noted the actual reduction of bandits as a whole, and, secondly, the shrinkage in size of bandit groups. In 1932 some bandit groups had an actual fighting force of 30,000 men, but at present the average is below 50 bandits per group. The chief cause for the existence of bandits in Manchoukuo is believed to be an economic one, resulting from unemployment."

STATISTICS OF BANDITRY

Size of bandit groups						
1 to 30	31 to 50	51 to 100	101 to 200	201 to 300	301 to 500	>501
Corresponding number of raids						
28,145	4,784	3,864	1,530	455	240	130

The first step towards making a comparison with larger fatal quarrels is to compensate for the unequal ranges of size in the above quotation, by dividing each number of raids by the corresponding range of size. For example $4784/(50-31+1)=239.2$. Any such ratio may be called the "number of raids per unit range of membership." For the first six groups it runs as follows

938.17 239.2 77.28 15.30 4.55 1.20

and shows plainly that *the smaller incidents were much the more frequent.*

Was a raid like a small battle or like a small war? In other words did the same group of bandits perpetrate many raids? This can be answered by attending to totals. The total number of appearances of bandits can be underestimated from the above table by multiplying the least size in each column by the corresponding number of raids, and comes to 758,371 or more. A moderate estimate is 1.3 million.

But the total number of bandits is stated to have been much less than 210,000. It is evident therefore that on the average the same bandit appeared in several raids. That is to say *a raid resembled a battle rather than a war.*

In the same year and same region the numbers killed in connection with banditry are stated thus: bandits 13,338, suppression troops 1361, civilians 2512. The total dead was therefore 17,211. This is 1.3 per cent of the total number of appearances of bandits.

The facts about banditry are plotted in Figure 2. This is not a $\phi(\mu)$ diagram, yet it is rather like one. For the abscissa, ν say, is $\log_{10}$ of the number of bandits in the group, whereas μ is $\log_{10}$ of the number of quarrel-dead. Again the ordinate, ψ say, is $\log_{10}$ of the average number of raids per unit range of membership, whereas ϕ is $\log_{10}$ of the average number of fatal quarrels per unit range of quarrel-dead. The first group in the data has a membership range of from 1 to 30, which is in a ratio too great for present purposes. Although the first group is shown on Figure 2, it will be ignored in the discussion. The blunted top of Figure 1 has however some resemblance to the horizontal in Figure 2. The other groups are well fitted by a straight line having the slope

$$d\psi(\nu)/d\nu = -2.29. \tag{7}$$

This has a remarkable resemblance to the mean slope, -2.38, across the gap in the world-diagram, as stated in (6). Can this agreement of the slopes be a mere coincidence? Or is it a clue to a general law concerning aggregation for aggression? It is at any rate easily explained by the following two assumptions:—

Firstly let us suppose that the quarrel-dead were on the average proportional to the number of bandits in the group. That is to say that groups continued to make raids until a constant fraction of their members had been killed. The supposition is that

$$n = A \times \text{(number of bandits in group)}$$

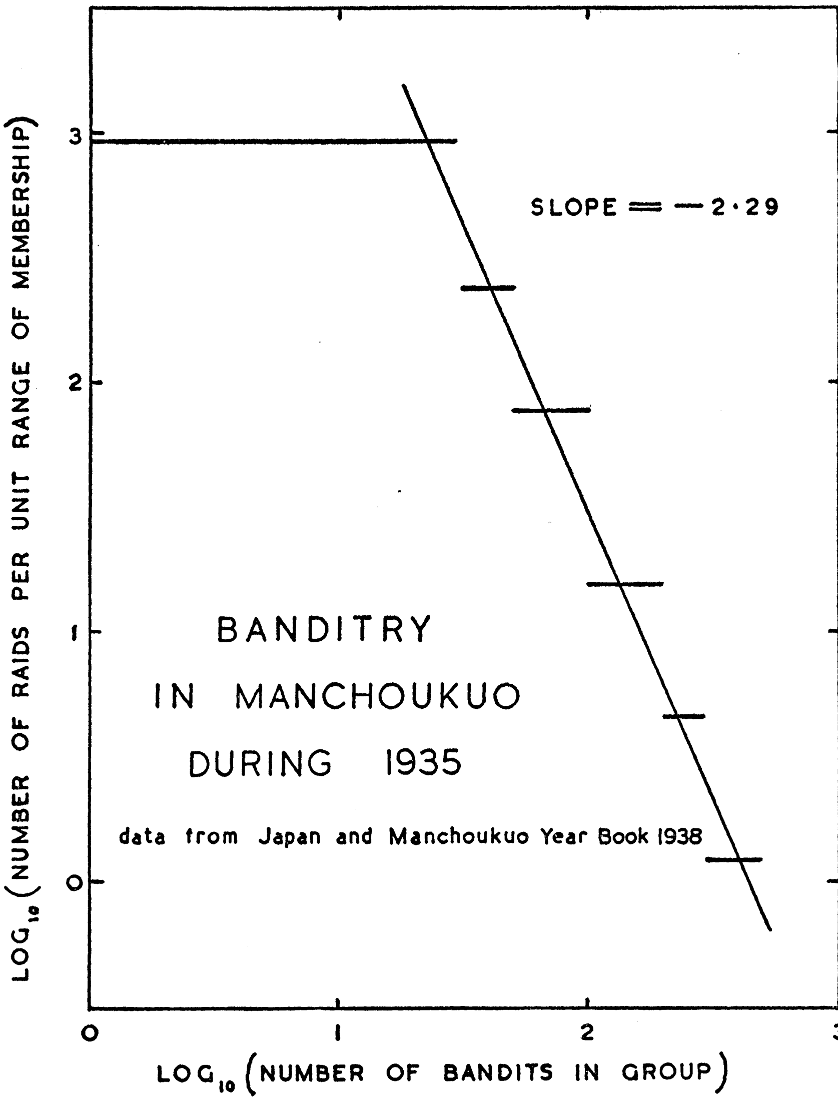

FIGURE 2

where A is independent of the size of the group. Accordingly

$$\mu = \nu + \log_{10} A. \tag{8}$$

It may not be necessary thus to suppose that the percentage of killed was exactly the same for every group, but only to suppose that it had no correlation with the size of the group.

Secondly let us suppose that the number of raids which a group of bandits made, as part of their quarrel with the rest of the community, was not correlated with the size of the group. More precisely the supposition is that:

$$\text{(number of raids made by a group)} = Bq(n)$$

where B is independent of n, so that

$$\psi(\nu) = \phi(\mu) + \log_{10} B. \tag{9}$$

According to these assumptions the $\psi(\nu)$ graph, when slid parallel to itself, horizontally through log A and vertically through log B, becomes a *local* $\phi(\mu)$ graph. The remarkable agreement of the slopes can therefore be interpreted as meaning that one of the characteristics of aggregation for aggression was the same in Manchoukuo as it was in the world-total.

That broad agreement is interesting; but there remain some doubts or discrepancies in detail. The Manchurian data do not fix A and B, though A can be roughly estimated. For it has been shown that on the average 1.3 per cent of the bandits who appeared in a raid were killed, and that a group on the average perpetrated several raids. The total quarrel-dead, n, may have been about 10 per cent of the number of bandits in the group; that is to say $A=0.1$, so that $\mu=\nu-1$. The mean value of μ on the straight part of the bandit-diagram is accordingly about one. If so, there is a discrepancy with the circular arc on the world-diagram. For at $\mu=1$ the slope of the arc is decidedly steeper, namely $d\phi/d\mu=-3.1$. The Manchurian data therefore suggest that the gap in the world-diagram (Figure 1) should be closed, not by the circular arc, but by the straight segment of slope -2.38 leaving discontinuities of slope at its two ends. I regard that as a suggestion to be remembered, but not to be acted on without further evidence. The shift $\mu=\nu-1$ cuts off most of the horizontal part of Figure 2 and thus increases its resemblance to the top of Figure 1.

GANGING IN CHICAGO

F. M. Thrasher [16] made a study of 1313 gangs in Chicago with a view to their re-direction, or to some other social improvement.

He defined a gang in these words (p. 57): "The gang is an interstitial group originally formed spontaneously, and then integrated through conflict. It is characterized by the following types of behaviour: meeting face to face, milling, movement through space as a unit, conflict, and planning. The result of this collective behavior is the development of tradition, unreflective internal structure, *esprit de corps*, solidarity, morale, group awareness, and attachment to a local territory."

Thrasher gives (p. 319) a table showing the approximate numbers of members in 895 gangs. The first two columns of the following table are copied from Thrasher, the third column is a deduction, designed to smooth out the disparities of the given ranges. The number of members

Number of members in gang	Number of such gangs	Gangs per unit range of membership
inclusively		
3 to 5	37	12.3
6 to 10	198	39.6
11 to 15	191	38.2
16 to 20	149	29.8
21 to 25	79	15.8
26 to 30	46	9.2
31 to 40	55	5.5
41 to 50	51	5.1
51 to 75	26	1.04
76 to 100	25	1.00
101 to 200	25	0.25
201 to 500	11	0.37
501 to 2,000	2	0.0013

in a gang is the same sort of quantity as the number of bandits in a group; so the same symbol ν is here used for $\log_{10}$ of the first column. Log_{10} of the third column is here denoted by $\chi(\nu)$ to distinguish it from $\psi(\nu)$, because the number of gangs is not similar to the number of raids, though related to it. Figure 3 shows a graph of $\chi(\nu)$. Apart from gangs of 15 or less, which show again the blunted top, the $\chi(\nu)$ graph is well fitted by the straight line shown, which has a slope

$$d\chi(\nu)/d\nu = -2.30. \tag{10}$$

It is remarkable that this slope is almost the same as $d\psi(\nu)/d\nu = -2.29$ for banditry in Manchoukuo. This agreement strengthens the suspicion that some fairly general tendency concerning aggregation for aggression is revealed by these otherwise scattered phenomena. Hypothetical explanation is again easy. The distinction between the number of gangs and the number of raids would not affect the slope, if it were true that the number, C say, of raids made by a gang was independent of the size of the gang. For then

$$\psi(\nu) = \chi(\nu) + \log_{10} C \tag{11}$$

and the constant $\log_{10} C$ would disappear on taking $d/d\nu$. Yet several features remain unclear. The Chicago gangs were certainly aggressive.

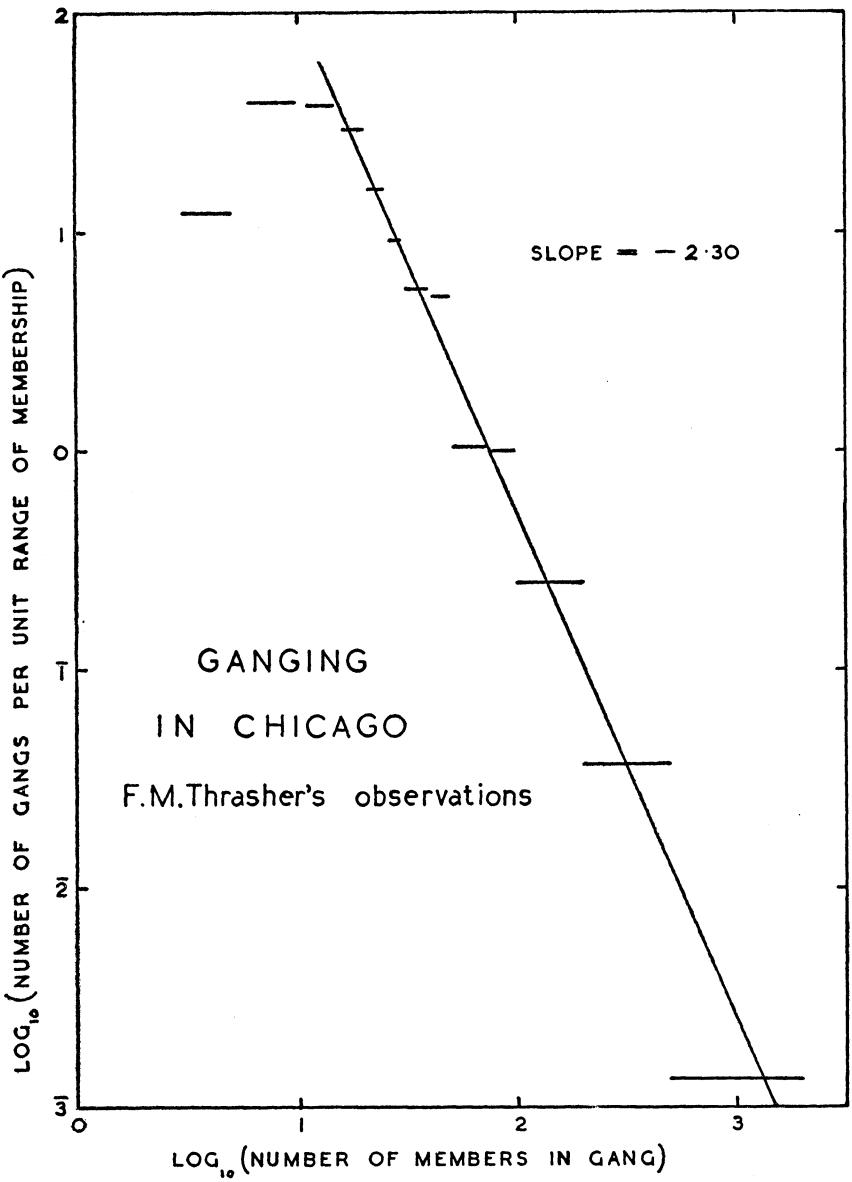

FIGURE 3

Thrasher has a chapter headed "Gang Warfare" and beginning with these words

"The gang is a conflict group. It develops through strife and thrives on war-

> fare. The members of a gang will fight each other. They will even fight for a 'cause.' . . . Gangsters are impelled, in a way, to fight; so much of their activity is outside the law that fighting is the only means of avenging injuries and maintaining the code."

Thrasher mentioned a few gang-fights which went as far as homicide, but he did not give any comprehensive statistics of the quarrel-dead. I am indebted to Prof. Lundberg and Margaret Black Richardson, for the indication of Thrasher's work.

SUMMARY

The $\phi(\mu)$ diagram in Figure 1 is the best conspectus of the facts for the world as a whole. Alternative tracks are shown across the gap where world-totals are lacking. Samples from Manchoukuo and Chicago confirm the slope of the straight alternative, whereas continuity of slope is a consideration in favor of the circular arc.

MORE EXACT CONNECTION BETWEEN THE $\phi(\mu)$ GRAPH OF FIGURE 1 AND THE NUMBERS OF FATAL QUARRELS IN UNIT RANGES OF MAGNITUDE. INTERPOLATION ACROSS THE GAP.

We need to connect the mean over a range with the value at its mid-point. It might be thought that the connection was obvious; for a glance at Figure 1 shows that the graph is fairly straight; and of course the mean ordinate of a straight segment is simply the ordinate at its mid-point. But that would be a delusion based on a doubly wrong type of mean. For here the mean of q, not of the ordinate $\log_{10} q$, has to be taken over the integers in the range of n, and not over the abscissa $\mu = \log_{10} n$. It will be shown that the logarithmic transformation of both coordinates introduces correcting factors which depend on the slope, vanishing when it is either 0 or -2, but otherwise often considerable. They are obtained from the simplest formula that can fit the facts tolerably, namely the formula for a straight line on the $\phi(\mu)$ graph. The straight lines are suitably short segments; for it would be a mistake to smooth away significant detail by the over-wide sweep of any too-simple formula. In particular the sharp bend near the top of Figure 1 comes from British criminal statistics where it appears to be significant; and something rather like it is to be seen in the graphs which represent banditry in Manchoukuo and ganging in Chicago; so this bend should certainly not be smoothed away. But elsewhere a short length of the $\phi(\mu)$ graph can be represented by

$$\phi(\mu) = B - C\mu$$

where B and C are positive constants. On taking antilogarithms this becomes

$$q = 10^B n^{-c}$$

an inverse power law. To find $\bar{q}$ this n^{-c} has to be averaged over a range of n. As n proceeds only by integers, the average should be made by sums, not by integrals. However the integral $\int n^{-c}dn$ gives a useful approximation to the sum.

This programme will now be worked out in proper detail.

The notation is that already defined in connection with the graph. The typical range ends, say, at $\mu = \mu_0 \pm \frac{1}{2}$, where $\mu_0 = 7, 6, 5, 4, 3, 2, 1$. It is convenient to begin with a rough and ready approximation, R, to the number Q of fatal quarrels in the range, whereby $\bar{q}(n)$ at the midpoint is multiplied by the number of integers in the range of quarrel-dead, which it is convenient for a later purpose, to denote by $\sum_n 1$ where $\sum_n$ sums for all n in the range of μ. Accordingly

$$R = 10^{\phi(\mu_0)} \sum_n 1. \tag{12}$$

For example, as read from the circular arc on Figure 1:

Range of magnitude $\mu = \log_{10} n$	Ordinate at midpoint $\phi(\mu_0)$	$\therefore \bar{q}(n)$ at midpoint	Number of integers in the range of quarrel-dead $\sum_n 1$	First approximation to number of quarrels R
0.5 to 1.5	3.87	7410	28	207,000
1.5 to 2.5	1.23	17.0	285	4,850
2.5 to 3.5	$\bar{1}.17$	0.131	2846	373

The above easy method is however suspect, because $\bar{q}(n)$ varies in the ratio 1740 in one of the ranges thus represented by its midpoint. So we must consider corrections.

Correction for width of range

As the $\phi(\mu)$ graph is nearly or quite straight, let us represent a local portion of it by

$$\phi(\mu) = \phi(\mu_0) - c(\mu - \mu_0) \tag{13}$$

where μ_0 is the midpoint of the range and $c = -d\phi/d\mu$ there. The distinction between $q(n)$ and $\bar{q}(n)$ is one that concerns detailed observations. It does not arise when, as now, deductions from a smooth curve in question. From (13)

$$q(n) = 10^{\phi(\mu)} = 10^{\phi(\mu_0)+c\mu_0}n^{-c}. \tag{14}$$

Let Q denote the total number of fatal quarrels in the range, then

$$Q = \sum_n q(n) = 10^{\phi(\mu_0)+c\mu_0} \sum_n n^{-c} \tag{15}$$

So from (12) and (15)

$$Q = R 10^{c\mu_0} \sum_n n^{-c} \Big/ \sum_n 1. \tag{16}$$

The coefficient of R will be called the correcting factor. Two verifications of (16) may be noted: (i) if $c=0$, then $Q=R$. (ii) if the range of quarrel-dead contained only a single integer n, and if it were located at μ_0 so that $10^{\mu_0}=n$, then (16) would become $Q=Rn^c n^{-c}/1=R$, as it should.

The sum in (16) was evaluated by a method connected with the "integral test" for the convergence of series, in the following manner. Let

$$b = 10^{\mu_0+1/2} \quad \text{and} \quad a = 10^{\mu_0-1/2} \text{ for brevity.} \tag{17}$$

In the following chain of equations, the first is an approximation which improves as μ_0 increases, and the other are accurate.

$$\sum_n n^{-c} \simeq \int_a^b t^{-c}dt = \frac{1}{1-c}(b^{1-c} - a^{1-c})$$
$$= \frac{10^{\mu_0(1-c)}}{1-c}\left\{10^{(1-c)/2} - 10^{(c-1)/2}\right\}. \tag{18}$$

In the special case of $c=0$, (18) gives the approximation

$$\sum_n 1 = 10^{\mu_0}(10^{1/2} - 10^{-1/2}) = b - a. \tag{19}$$

Strictly $\sum_n 1$ is the integer next below $b-a$; but it is advisable to use the same type of approximation in both numerator and denominator of (16). Insertion of (18) and (19) into (16) gives the correcting factor

$$Q/R = \frac{1}{1-c}\,\frac{10^{(1-c)/2} - 10^{(c-1)/2}}{10^{1/2} - 10^{-1/2}}. \tag{20}$$

This expression is unity when $c=0$ or $c=2$.
For the wars, $c=1.50$ and $Q/R=0.855$.

The observed quantity is Q in the unit range of magnitude. The $\phi(\mu)$ curve should however be connected to μ at a point, and so to R. We have

$$\log_{10} R = \log_{10} Q - \log_{10} 0.855 = \log_{10} Q + 0.068.$$

In the $\phi(\mu)$ graph (Figure 1) the sloping line for the wars was therefore raised 0.068 above that which best fits the observations grouped in unit ranges of magnitude.

In the unexplored region we have, on taking the first approximation R from the previous table:

World-totals for the 126 years 1820 to 1945

Range of magnitude μ	$c = -\frac{d\phi}{d\mu}$	Correcting factor Q/R	Corrected number of fatal quarrels Q	Range of number killed in a quarrel n
0.5 to 1.5	3.14	1.92	397,000	4 to 31
1.5 to 2.5	2.32	1.16	5,630	32 to 316
2.5 to 3.5	1.87	0.95	354	317 to 3162

The fourth column shows the improved estimates for the unexplored region between the wars and the murders according to the circular arc in Figure (1). The difficulties of direct counting have already been described, and they are emphasized by these large numbers. In the range where 354 are expected, I have counted 188. If the straight line in Figure (1) were accepted instead of the circular arc, 354 would be increased to 2530. Although I know that the search is incomplete, I am unable to believe that less than a tenth have been found; and so I prefer the circular arc.

THE TOTAL NUMBER OF PERSONS WHO DIED BECAUSE OF QUARRELS DURING THE 126 YEARS FROM 1820 TO 1945 A.D.

Ends of range of magnitude	Total number of deaths in millions	How computed
$7\pm\frac{1}{2}$	36	By summation over a list of fatal quarrels.
$6\pm\frac{1}{2}$	6.7	
$5\pm\frac{1}{2}$	3.4	
$4\pm\frac{1}{2}$	0.75	
$3\pm\frac{1}{2}$	0.30	From the circular arc in Figure (1) by the method described below.
$2\pm\frac{1}{2}$	0.40	
$1\pm\frac{1}{2}$	2.2	
$0\pm\frac{1}{2}$	9.7	From page 528
	Total 59	

A remarkable feature of the above table is that the heavy loss of life occurred at the two ends of the sequence of magnitudes, namely the World Wars and the murders. The small wars contributed much less to the total. The total deaths because of quarrels should be compared with the *total deaths from all causes.* There are particulars given by de Jastrzebski [9], and in the Statistical Year Books of the League of Nations, which allow the total to be estimated. A mean world population of 1.5×10^9 and a mean death rate of 20 per thousand per year would give during 126 years 3.8×10^9 deaths from all causes. Of these the part caused by quarrels was 1.6 per cent. This is less than one might have guessed from the large amount of attention which quarrels attract. Those who enjoy wars can excuse their taste by saying that wars after all are much less deadly than disease.

The method of computation in the unexplored region was an extension of that explained above. The number D of dead in any range of μ is strictly

$$D = \sum_n nq(n) = \sum_n 10^{\mu+\phi(\mu)}. \tag{21}$$

First a common-sense estimate, E of D was obtained from the graph at the midpoint μ_0 of the range $\mu_0 - \frac{1}{2} \leqq \mu \leqq \mu_0 + \frac{1}{2}$ thus, by (12),

$$E = 10^{\mu_0}R = 10^{\mu_0+\phi(\mu_0)} \sum_n 1. \tag{22}$$

Then E was corrected by multiplication by a factor which was found to be approximately

$$\frac{D}{E} \doteq \frac{1}{2-c} \left\{ \frac{10^{(2-c)/2} - 10^{(c-2)/2}}{10^{1/2} - 10^{-1/2}} \right\}. \tag{23}$$

The theory of this correction is based on the approximation (13), whence it follows that

$$\mu + \phi(\mu) = c\mu_0 + \phi_0 + \mu(1-c). \tag{24}$$

From (21) and (24)

$$D = 10^{c\mu_0+\phi(\mu_0)} \sum_n n^{1-c}, \quad \text{and} \tag{25}$$

$$\frac{D}{E} = 10^{\mu_0(c-1)} \frac{\sum_n n^{1-c}}{\sum_n 1}. \tag{26}$$

The sums in (26) were obtained from (18) by suitable alterations of the index and they lead to (22). The correction (22) was verified in particular cases by the "deferred approach to the limit" of Richardson and Gaunt [14].

COMPARISON WITH THE DISTRIBUTION OF SIZES OF TOWNS

Towns and wars are both examples of human aggregation. A referee has asked me to compare them. For simplicity I take the distribution of towns from Lotka's book [10A] in Auerbach's idealized form whereby the town of rank r in a given country has a population n such that

$$nr = A, \tag{1}$$

in which A is a constant, namely the population of the largest town. The desired comparison might conceivably be made by arranging the fatal quarrels in order of rank; but there would be many artificially coincident ranks on account of the rounding off of imperfectly known casualties; moreover the gap in the data prevents the extension of rank from the wars to the murders. It is preferable therefore to leave the fatal quarrels as they are already shown on Figure 1, and to transform Auerbach's law so as to see how towns would appear on a diagram of that type. For the purpose of comparison the same symbols will be used in corresponding meanings. Thus in this section

$$n \text{ is the population of a town, and } \mu \text{ will be } \log_{10} n. \tag{2}$$

Again, to match the fatal quarrels, $\bar{q}$ will here be defined as

$$\bar{q} = \frac{\text{number of towns in a range of population}}{\text{number of integers in that range of population}} \cdot \tag{3}$$

Also, as before, let

$$\phi = \log_{10} \bar{q}. \tag{4}$$

Consider two towns of ranks r', r'', of which r'' is the greater. Let them mark the ends of a range. The number of towns inside the range is strictly $r''-r'-1$. But let us follow the usual statistical practice of regarding half of an end-object as lying on either side of the end. The number of towns in the range is then simply $r''-r'$. Let n' and n'' be the respective populations. With the same convention about ends on the population range, we have from (3)

$$\bar{q} = \frac{r'' - r'}{n' - n''}. \tag{5}$$

Elimination of ranks between (5) and (1) gives

$$\bar{q} = \frac{A}{n'n''}. \tag{6}$$

So that from (4)

$$\phi = \log_{10} A - \log_{10} n' - \log_{10} n''. \tag{7}$$

As for the wars, let μ_0 be the midpoint of the range of μ. It is

$$\mu_0 = \tfrac{1}{2}(\log_{10} n' + \log_{10} n''). \tag{8}$$

From (7) and (8)

$$\phi = \log_{10} A - 2\mu_0, \tag{9}$$

however wide or narrow the range may be. That the constant $\log_{10} A$ should be the same whatever the grouping is a peculiarity of the slope

$$d\phi/d\mu_0 = -2 \tag{10}$$

as shown on page 541, where the correction vanished for $c = 2$.

This slope for the towns is quite close to that of the straight line which in Figure 1 might join the murders to the world wars. Such a broad resemblance between two forms of aggregation is certainly interesting, and may suggest theories. Perhaps the most suggestive formula is

$$q = An^{-2} \tag{11}$$

which follows from (6) when n' is almost equal to n''. Rashevsky [11A], who gives formula (11), has tried-out several explanations.

These overall resemblances do not however conduce to an accurate description of fatal quarrels; for they distract attention from the curvature of the $\phi(\mu)$ graph in Figure 1. The slope of the part relating to wars is certainly not -2. If it were so, then the number of fatal quarrels would increase 10 times for each unit decrease of magnitude. The observed ratio is less than four.

REFERENCES

[1] Anderson, J. G. 1939. *China fights for the world.* Kegan Paul, London, p. 106.
[1A] Carr-Saunders, A. M. 1936. *World Population.* Clarendon Press, Oxford.
[2] Coupland, R. 1939. *The Exploitation of East Africa 1856–1890.* Faber & Faber, London. Chh. VII, XI.
[3] Durbin, E. F. M., and Bowlby, J. 1939. *Personal Aggressiveness and War.* Kegan Paul, London.
[4] Fisher, R. A. 1922. *Phil. Trans. Roy. Soc. Lond.* A **222**, 332.
[5] Glover, E. 1933 & 1947. *War, Sadism & Pacifism.* Allen & Unwin, London.
[6] Harding, D. W. 1941. *The Impulse to Dominate.* Allen & Unwin, London.
[7] Hume, M. A. S. 1900. *Modern Spain.* Fisher Unwin, London.
[8] *Japan and Manchoukuo Year Book,* 1938. Tokio.
[9] Jastrzebski, S. de. 1929 article *Death-Rate* in Ency. Brit. XIV edn.
[10] Jeffreys, H. 1939. *Theory of Probability,* p. 148. Clarendon Press, Oxford.
[10A] Lotka, A. J. 1925. *Elements of Physical Biology.* Williams & Wilkins, Baltimore.
[11] Miliukov, P. article *Russia* in *Ency. Brit.* XIV ed., **19**, 720 d.
[11A] Rashevsky, N. 1947. *Mathematical Theory of Human Relations.* The Principia Press, Bloomington, Indiana.
[12] Richardson, L. F. in *Nature* of 15 Nov. 1941.
[13] Richardson, L. F. 1945. *J. Roy. Statistical Soc.* **107,** 242–250.
[14] Richardson, L. F., and Gaunt, J. A. 1927. *Phil. Trans. Roy. Soc. Lond.* A **226,** 299–361.
[15] Sleeman, W. H. 1836. *Ramaseeana.* Military Orphan Press Calcutta, p. 39.
[16] Thrasher, F. M. 1927. *The Gang.* Chicago University Press.
[17] Whitaker's Almanack, annual, London.
[18] Wright, Q. 1942. *A Study of War.* Chicago University Press.