J. R. Statist. Soc. A,
(1975), **138**, *Part* 2, p. 204

The Generalized Waring Distribution. Part II

By J. O. Irwin

(*Retired*)

Summary

Part II is aimed at exploring the long-tailed G.W.D. distributions. These have been tabulated on an electronic computer for $\rho = 0{\cdot}5\ (1{\cdot}0)\ 4{\cdot}5$ and $q_a, q_k = 0{\cdot}25, 0{\cdot}5\ (0{\cdot}1)\ 0{\cdot}9$. Certain negative binomials, which are limiting forms of the G.W. distributions when q_a/p_a is finite, $\rho \to \infty$, $a \to \infty$, have been tabulated with the above values of q_a and $k = 4{\cdot}5q_k/p_k$ [$q_k = 0{\cdot}25$, $0{\cdot}5\ (0{\cdot}1)\ 0{\cdot}9$]. The upper 10, 5, 1, 0·1, 0·01 and 0·005 per cent points were recorded on the computer sheets. The distributions were not taken to more than 2,000 terms in the computer calculations. In some cases the percentage points were much greater than 2,000. Here it was, as a rule, possible to obtain the percentage points by extrapolation on a desk machine. The mode, median and mean were obtained for each distribution. So were the standard deviation and coefficient of variation. The successive sub-sections of Part II deal (i) with the mode, median and mean and their inter-relations, (ii) with the standard deviation and coefficient of variation, (iii) with the upper percentage points of the tabulated G.W. distributions (ρ finite), (iv) with the same measures of location and dispersion, and the same percentage points for the tabulated negative binomials. The interrelation of all these quantities for different values of the parameters throws much light on the form of the G.W. distribution over all possible values of the parameters a, k, ρ. The standard deviation and coefficient of variation are not very appropriate measures of scale or dispersion for $\rho \leqslant 4$. Alternatives are considered. Also the generally used measures of skewness are inappropriate or inapplicable in such cases; "length of tail" defined by the ratio of suitably chosen percentage points seems more appropriate. However, the standard deviation has one desirable property. The percentage points in terms of the S.D. for fixed ρ and varying q_a, q_k are almost independent of the two latter parameters. This section ends with an illustrative example.

Keywords: COMPUTER TABULATION; MEAN, MEDIAN AND MODE; MEASURES OF DISPERSION; PERCENTAGE POINTS; SKEWNESS; LENGTH OF TAIL

6. The Electronic Computer Tabulations†

6.1. One of the main objects of this study was to explore the long-tailed G.W. distributions. Dr David Hill, to whom I am greatly indebted, tabulated these for $\rho = 0{\cdot}5\ (1{\cdot}0)\ 4{\cdot}5$ and $q_a, q_k = 0{\cdot}25, 0{\cdot}5\ (0{\cdot}1)\ 0{\cdot}9$. Certain negative binomial distributions were also tabulated to compare this limiting form with the corresponding G.W. distributions for $\rho = 4{\cdot}5$ and the same values of q_a, q_k. The negative binomial distributions selected were $\{(1-q_a A)/p_a\}^{-k}$. These are the limiting forms of the G.W. distributions when $\rho \to \infty$ (and consequently $a = \rho q_a/p_a \to \infty$) while k was given the values $4{\cdot}5\, q_k/p_k$, with $q_k = 0{\cdot}25, 0{\cdot}5\ (0{\cdot}1)\ 0{\cdot}9$ (see end of p. 220). These were printed for $x = 0\ (1)\ 99\ (5)\ 499\ (20)\ 999$. After the 1,000th term every 100th was printed and

† The numbering of Sections continues from Part I.

the tabulations were taken to 2,000 terms, unless the *upper* 0·005 per cent point (corresponding to the probability 0·00005 in the upper tail) was reached first. The upper 10, 5, 1, 0·1, 0·01 and 0·005 per cent points were also recorded on the computer sheets, whenever these were reached in 2,000 terms. In a good many cases the higher percentage points (even the 5 per cent point for $\rho = 0{\cdot}5$, $q_a = q_k = 0{\cdot}9$) were much greater than 2,000; but, in most cases, I have been able to extrapolate for their approximate values on a desk machine. Great accuracy is not required, but it is of interest to know their order of magnitude. When $\rho = 0{\cdot}5$ the extrapolation was not possible for the 0·1, 0·01 and 0·005 per cent points. The tails are here so long that, for example, the 1 per cent point for the distribution with $\rho = 0{\cdot}5$, $q_a = q_k = 0{\cdot}9$ is approximately 238,000. (See Section 6.7 for further details.)

Dr Hill also tabulated the mode while the median was obtainable at sight from his results. The mean $ak(\rho-1)$ is easily tabulated. (For $\rho = 0{\cdot}5$, the mean is infinite.) I am grateful to Miss Irene Allen for extracting the information I required from Dr Hill's voluminous results. The final form of the tables in this paper is, however, my own.

6.2. The mode, median and mean are given in Table 1; the standard deviation and coefficient of variation in Table 2; while Table 3 gives the percentage points. The results are considered in some detail below; however, a general remark about the negative binomial distributions tabulated may be made here.

The negative binomials were selected to see how much difference is made in the form of the G.W. distributions (defined by given ρ, q_a, q_k) when ρ increases from 4·5 to infinity, q_a and $k = 4{\cdot}5q_k/p_k$ remaining fixed. Of course, as ρ increases from 4·5 to infinity in this way, $q_k = k/(k+\rho)$ does not remain fixed; $q_k \to 0$ as $\rho \to \infty$. The difference between any particular G.W. distribution and this limiting form becomes more striking as ρ diminishes. A complete comparison would have necessitated the calculation of the negative binomials corresponding to the G.W. distributions with $\rho = 0{\cdot}5, 1{\cdot}5, 2{\cdot}5, 3{\cdot}5$, as well as $\rho = 4{\cdot}5$. This was not done. However, two comparisons of the kind may be made from the tables. When $\rho = 4{\cdot}5$ and q_k takes the values 0·25, 0·5, 0·6, 0·7 0·8, 0·9, $k = 1{\cdot}5, 4{\cdot}5, 6{\cdot}75, 10{\cdot}5, 18, 40{\cdot}5$. But $k = 1{\cdot}5$ also when $\rho = 1{\cdot}5$, $q_k = 0{\cdot}5$; $k = 4{\cdot}5$ also when $\rho = 0{\cdot}5$, $q_k = 0{\cdot}9$. Other comparisons of this kind would require interpolation in the tables.

In each block of Tables 1 and 2, corresponding to fixed but finite values of ρ, the entries must of course be symmetrical in q_a, q_k. For $\rho = 0{\cdot}5, 1{\cdot}5, 2{\cdot}5, 3{\cdot}5$; the entries for $q_k \geqslant q_a$ are given; equal values are not repeated twice. In the negative binomials, however, the entries are in general not symmetrical; but the mean is an exception to this. For example, if $q_a = 0{\cdot}5$, $k = 4{\cdot}5q_k/p_k$ where q_k and p_k refer to the G.W. distribution with $\rho = 4{\cdot}5$—not to the negative binomial for which $q_k \to 0$ when $\rho \to \infty$—we have, for $q_k = 0{\cdot}9$, the negative binomial $(2-A)^{-40{\cdot}5}$; but, if $q_a = 0{\cdot}9$, $q_k = 0{\cdot}5$, we have $(10-9A)^{-4{\cdot}5}$. Nevertheless, the means of the two distributions are the same, namely 40·5. The most interesting comparison involving the negative binomials is the column by column (or row by row) comparison between them (section (vi) of the tables) and the G.W. distributions for $\rho = 4{\cdot}5$ (section (v)). For this reason the pairs of equal entries are both given for $\rho = 4{\cdot}5$.

Table 3 is arranged slightly differently from Tables 1 and 2; so that the percentage points for the same distribution occur together in a column. The symmetry for any fixed percentage point and fixed ρ, when q_a and q_k are interchanged, still exists; but as the equal values are more widely separated, they have been repeated in their appropriate places. This facilitates comparison of the values of a fixed percentage point for fixed ρ and q_a for varying q_k, fixed ρ and q_k for varying q_a and fixed q_a

TABLE 1

Mode, median and mean for long-tailed G.W. distribution

	Mode						*Median*						*Mean*					
(i) $\rho = 0\cdot5$																		
	q_k						q_k						q_k					
q_a	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9
0·25	0	0	0	0	0	0	0	0	0	0	0	0						
0·5	—	0	0	0	0	0	—	0	0	1	1	4						
0·6	—	—	0	0	0	0	—	—	1	1	3	8	(Mean is infinite)					
0·7	—	—	—	0	0	0	—	—	—	3	6	15						
0·8	—	—	—	—	0	2	—	—	—	—	12	30						
0·9	—	—	—	—	—	8	—	—	—	—	—	75						
(ii) $\rho = 1\cdot5$																		
	q_k						q_k						q_k					
q_a	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9
0·25	0	0	0	0	0	0	0	0	0	0	1	2	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{4}$	$3\frac{1}{2}$	6	$13\frac{1}{2}$
0·5	—	0	0	0	1	2	—	1	2	3	5	13	—	$4\frac{1}{2}$	$6\frac{3}{4}$	$10\frac{1}{2}$	18	$40\frac{1}{2}$
0·6	—	—	0	1	2	6	—	—	3	5	9	21	—	—	$10\frac{1}{8}$	$15\frac{3}{4}$	27	$60\frac{3}{4}$
0·7	—	—	—	2	5	12	—	—	—	8	15	35	—	—	—	$25\frac{1}{2}$	42	$94\frac{1}{2}$
0·8	—	—	—	—	10	25	—	—	—	—	27	63	—	—	—	—	72	162
0·9	—	—	—	—	—	62	—	—	—	—	—	146	—	—	—	—	—	$364\frac{1}{2}$
(iii) $\rho = 2\cdot5$																		
	q_k						q_k						q_k					
q_a	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9
0·25	0	0	0	0	0	0	0	0	1	1	2	5	$\frac{25}{54}$	$1\frac{7}{18}$	$2\frac{1}{12}$	$3\frac{13}{54}$	$5\frac{5}{9}$	$12\frac{1}{2}$
0·5	—	0	1	2	3	9	—	2	3	5	9	22	—	$4\frac{1}{6}$	$6\frac{1}{4}$	$9\frac{13}{18}$	$16\frac{2}{3}$	$37\frac{1}{2}$
0·6	—	—	2	3	7	16	—	—	5	8	15	35	—	—	$9\frac{3}{8}$	$14\frac{7}{12}$	25	$56\frac{1}{4}$
0·7	—	—	—	6	12	29	—	—	—	14	24	56	—	—	—	$22\frac{37}{54}$	$38\frac{8}{9}$	$87\frac{1}{2}$
0·8	—	—	—	—	23	55	—	—	—	—	43	98	—	—	—	—	$66\frac{2}{3}$	150
0·9	—	—	—	—	—	132	—	—	—	—	—	226	—	—	—	—	—	$337\frac{1}{2}$

TABLE 1 (*cont.*)

(iv) $\rho = 3{\cdot}5$

	q_k																	
q_a	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9
0·25	0	0	0	0	0	1	0	1	1	2	4	8	$\frac{49}{90}$	$1\frac{19}{30}$	$2\frac{19}{20}$	$3\frac{73}{90}$	$6\frac{8}{15}$	$14\frac{7}{10}$
0·5	—	1	2	3	7	17	—	3	5	8	18	31	—	$4\frac{9}{10}$	$7\frac{7}{20}$	$11\frac{13}{30}$	$19\frac{3}{5}$	$44\frac{1}{10}$
0·6	—	—	4	6	12	28	—	—	7	12	21	48	—	—	$11\frac{1}{40}$	$17\frac{3}{20}$	$29\frac{2}{5}$	$66\frac{3}{20}$
0·7	—	—	—	11	20	48	—	—	—	19	34	77	—	—	—	$26\frac{61}{90}$	$45\frac{11}{15}$	$102\frac{9}{10}$
0·8	—	—	—	—	37	88	—	—	—	—	59	134	—	—	—	—	$78\frac{2}{5}$	$176\frac{2}{5}$
0·9	—	—	—	—	—	206	—	—	—	—	—	306	—	—	—	—	—	$396\frac{9}{10}$

(v) $\rho = 4{\cdot}5$

	q_k, k																	
	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9
q_a	1·5	4·5	6·75	10·5	18·0	40·5	1·5	4·5	6·75	10·5	18·0	40·5	1·5	4·5	6·75	10·5	18·0	40·5
0·25	0	0	0	0	1	3	0	1	2	3	5	11	$\frac{9}{14}$	$1\frac{13}{14}$	$2\frac{25}{28}$	$4\frac{1}{2}$	$7\frac{5}{7}$	$17\frac{5}{14}$
0·5	0	2	3	6	10	25	1	4	6	10	17	40	$1\frac{13}{14}$	$5\frac{11}{14}$	$8\frac{19}{28}$	$13\frac{1}{2}$	$23\frac{1}{7}$	$52\frac{1}{14}$
0·6	0	3	6	9	17	41	2	6	10	15	27	62	$2\frac{25}{28}$	$8\frac{19}{28}$	$13\frac{1}{56}$	$20\frac{1}{4}$	$34\frac{5}{7}$	$78\frac{3}{28}$
0·7	0	6	9	16	29	68	3	10	15	25	43	98	$4\frac{1}{2}$	$13\frac{1}{2}$	$20\frac{1}{4}$	$31\frac{1}{2}$	54	$121\frac{1}{2}$
0·8	1	10	17	29	52	122	5	17	27	43	75	170	$7\frac{5}{7}$	$23\frac{1}{7}$	$34\frac{5}{7}$	54	$92\frac{5}{7}$	$208\frac{2}{7}$
0·9	3	25	41	68	122	283	11	40	62	98	170	387	$17\frac{5}{14}$	$52\frac{1}{14}$	$78\frac{3}{28}$	$121\frac{1}{2}$	$208\frac{2}{7}$	$468\frac{9}{14}$

(vi) Negative Binomials $\{(1-q_a A)/p_a\}^{-k}$, $k = 4{\cdot}5q_k/p_k$*

	q_k, k																	
	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9
q_a	1·5	4·5	6·75	10·5	18·0	40·5	1·5	4·5	6·75	10·5	18·0	40·5	1·5	4·5	6·75	10·5	18·0	40·5
0·25	0	1	1	3	5	13	0	1	2	3	6	13	$\frac{1}{2}$	$1\frac{1}{2}$	$2\frac{1}{4}$	$3\frac{1}{2}$	6	$13\frac{1}{2}$
0·5	0	3	5	9	17	39	1	4	6	10	17	40	$1\frac{1}{2}$	$4\frac{1}{2}$	$6\frac{3}{4}$	$10\frac{1}{2}$	18	$40\frac{1}{2}$
0·6	0	5	8	14	25	59	2	6	9	15	26	60	$2\frac{1}{4}$	$6\frac{3}{4}$	$10\frac{1}{8}$	$15\frac{3}{4}$	27	$60\frac{3}{4}$
0·7	1	8	13	22	39	93	3	10	15	24	41	94	$3\frac{1}{2}$	$10\frac{1}{2}$	$15\frac{3}{4}$	$24\frac{1}{2}$	42	$94\frac{1}{2}$
0·8	2	14	23	38	68	158	5	17	26	41	71	161	6	18	27	42	72	162
0·9	4	31	51	85	153	355	10	37	58	91	159	361	$13\frac{1}{2}$	$40\frac{1}{2}$	$60\frac{3}{4}$	$94\frac{1}{2}$	162	$364\frac{1}{2}$

* In the negative binomials $\rho \to \infty, a \to \infty, q_k \to 0$. The values of q_k in the *headings* are the same as in section (v) and k has been *chosen* to be $4{\cdot}5q_k/p_k$.

and q_k for varying ρ; for one has only to look down a column or across a row, at the entries which are in the same position in each block.

6 (CONT.). THE MODE, MEDIAN AND MEAN

6.3. Table 1 gives the Mode, Median and Mean. The Mode is the value of the variate $x = 0, 1, 2, 3, \ldots$ for which the frequency is greatest; if two frequencies are equal the greater value is tabulated. The Median tabulated is the first value of x for which the cumulative frequency is greater than 0·5. The Mean $ak/(\rho-1)$ is given *exactly* (as an integer + a proper fraction). This departure from custom is justified, because the proper fractions are all very simple. If two or three decimals had been given there would have been some loss of information and no saving of space.

The general properties of the mode have already been discussed. For fixed ρ, Table 1 shows clearly the changes in the mode, median and mean with q_a and q_k. It also illustrates the rule that for fixed q_a, q_k and unit increase in ρ, the mode changes by about $q_a q_k / p_a p_k$.

However, the most striking feature of the table arises from the comparison of the values of the mode, median and mean. The rule that (median − mode) = 2(mean − median) approximately, does not apply for such long-tailed distributions as these. Since for $\rho \leqslant 1$ the mean is infinite, one would expect the mode and median to be relatively close together, compared with the mean, for small values of $\rho > 1$. (We should deduct $\frac{1}{2}$ from the tabulated values of the median, before making the comparison.) Even for $\rho = 4{\cdot}5$, this is a very marked characteristic of the distributions, and still more so for $\rho = 3{\cdot}5$, 2·5 and 1·5.

When the first frequency of the distribution is more than 50 per cent, the mode and median are both tabulated as zero. In other cases, in comparing (median − mode) with (mean − median) we have said that 0·5 should first be deducted from the median. If r is the tabulated value, the "true" median could be any value between $(r-1)$ and r where $F(r) > \frac{1}{2} > F(r-1)$. For comparative purposes, it can be taken as $r - \frac{1}{2}$. With a strictly discrete distribution, this is the best that can be done. However, when r or $r - \frac{1}{2}$ is tabulated, the values do not run completely smoothly. Such values could only be obtained by a conventional interpolation based on a continuous curve. For the median, perhaps the histogram corresponding to the discrete distribution could be used.

It is clear that the rule, very nearly true for Pearson distributions, not heterotypic in K. Pearson's sense, that (median − mode) = 2(mean − median) fails completely here. If we write r_m = (median − mode)/(mean − median) we find instead of $r_m = 2$, the following approximate values:

$\rho = 1{\cdot}5$

q_a	q_k	r_m
0·25	0·25, 0·5 (0·1) 0·9	<0·1
0·5	0·5	0·1
⩾0·5	0·6 (0·1) 0·9	0·3–0·4

$\rho = 2{\cdot}5$

q_a	q_k	r_m
0·25	0·25, 0·5	<0·3
0·25	0·6, 0·7, 0·8	0·3
0·25	0·9	0·6
0·5	0·5	0·4
0·5	0·6, 0·7	0·6
0·5	0·8, 0·9	0·8
0·6	0·6, 0·7	0·6
0·6	0·8, 0·9	0·8
0·7	0·7, 0·8, 0·9	0·8
0·8, 0·9	0·8, 0·9	0·8

$\rho = 3{\cdot}5$			$\rho = 4{\cdot}5$		
q_a	q_k	r_m	q_a	q_k	r_m
0·25	0·25	<0·4	0·25	0·25	<0·4
0·25	0·5	0·4	0·25	0·5	0·4
0·25	0·6	0·4	0·25	0·6	1·1
0·25	0·7	0·5	0·25	0·7	1·1
0·25	0·8	1·0	0·25	0·8	1·1
0·25	0·9	1·0	0·25	0·9	1·1
0·5	0·5	0·6	0·5	0·5	0·7
0·5	0·6, 0·7, 0·8	0·9	0·5	0·6	0·8
0·5	0·9	1·0	0·5	0·7	0·9
			0·5	0·8	1·0
0·6	0·6, 0·7, 0·8	0·9	0·5	0·9	1·2
0·6	0·9	1·0			
			0·6	0·6, 0·7	1·0
0·7	0·7	0·9	0·6	0·8, 0·9	1·2
0·7	0·8, 0·9	1·1			
0·8, 0·9	0·8, 0·9	1·1	0·7	0·7, 0·8, 0·9	1·2
			0·8	0·8, 0·9	1·2
			0·9	0·9	1·3

In all cases q_a and q_k can, of course, be interchanged.

In the negative binomials of Table 1(vi) the mode is always greater than in the corresponding distribution of Table 1(v). The median is greater (by one or two units) only when $q_a = 0{\cdot}25$, $q_k = 0{\cdot}8$, $0{\cdot}9$ ($k = 18$, $40{\cdot}5$). Otherwise the median in section (vi) is at most equal to the corresponding median in section (v). In general the two are rather close.

The mean, median and mode are very close together in these negative binomials, compared with the G.W. distributions of (v); so that without some form of interpolation, perhaps unjustifiable, r_m cannot be estimated. The possible limits are too wide. However, for q_a, $q_k = 0{\cdot}8$ and $0{\cdot}9$, we can safely say that r_m is between 1 and 2.

It is interesting to compare the two sets of cases which give rise to the same value of k:

(a) (i) $\rho = 1{\cdot}5, q_k = 0{\cdot}5$, (ii) $\rho = 4{\cdot}5, q_k = 0{\cdot}25$, (iii) negative binomial with $k = 1{\cdot}5$;

(b) (i) $\rho = 0{\cdot}5, q_k = 0{\cdot}9$, (ii) $\rho = 4{\cdot}5, q_k = 0{\cdot}5$, (iii) negative binomial with $k = 4{\cdot}5$.

As (ii) and (iii) have already been compared, it remains only to compare (i) with the other cases.

The mode and median of (a) (i) and (a) (ii) are equal for all tabulated values of q_a except $q_a = 0{\cdot}9$, when (a) (i) has a mode of 2 and a median of 13; and (a) (ii) has a mode of 3 and a median of 11. In case (b), the mode of (b) (i) is small compared with the mode of (b) (ii) and the median of (b) (i) is for $q_k > 0{\cdot}5$ considerably larger than the median of (b) (ii). In (a) the decrease in ρ and the increase of q_k more or less compensate; in (b) the reduction of ρ to a value below unity altogether outweighs the increase of q_k.

6 (CONT.). THE STANDARD DEVIATION AND COEFFICIENT OF VARIATION

6.4. Table 2 gives the standard deviation and coefficient of variation for $\rho = 2{\cdot}5$, $3{\cdot}5$, $4{\cdot}5$. For $\rho = 0{\cdot}5$ and $1{\cdot}5$ (and in general for $\rho \leqslant 2$) the standard deviation is infinite. Thus it is not surprising that the values of the standard deviation in Table 2 increase rapidly (by 40–60 per cent) as ρ falls from 3·5 to 2·5. The difference between the values of the S.D. for $\rho = 4·5$ and $\rho = 3·5$ is much less. Indeed the values for $\rho = 4·5$, $q_a = 0·25$ and all q_k are a little greater than for $\rho = 3·5$.

TABLE 2

Standard deviation and coefficient of variation for long-tailed G.W. distribution

Standard deviation

	(iii) $\rho = 2·5$, q_k						(iv) $\rho = 3·5$, q_k					
q_a	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9
0·25	1·83	4·16	5·83	8·60	14·10	30·55	1·25	2·76	3·84	5·62	9·16	19·72
0·5	—	9·43	13·23	19·50	31·97	69·28	—	6·86	9·55	13·97	22·75	48·98
0·6	—	—	18·56	27·36	44·86	97·21	—	—	13·29	19·44	31·66	68·18
0·7	—	—	—	40·33	66·13	143·3	—	—	—	28·45	46·33	99·76
0·8	—	—	—	—	108·4	234·9	—	—	—	—	75·44	162·4
0·9	—	—	—	—	—	525·8	—	—	—	—	—	349·8

	(v) $\rho = 4·5$, q_k, k						(vi) Negative binomial $\{1/p_a - q_a A/p_a\}^{-k}$, $k(\rho \to \infty, a \to \infty)$					
	0·25	0·5	0·6	0·7	0·8	0·9						
q_a	1·5	4·5	6·75	10·5	18·0	40·5	1·5	4·5	6·75	10·5	18·0	40·5
0·25	1·36	2·97	4·12	6·00	9·74	20·89	0·82	1·41	1·73	2·16	2·83	4·25
0·5	2·97	6·57	9·02	13·15	21·33	45·77	1·73	3·00	3·67	4·58	6·00	9·00
0·6	4·12	9·02	12·50	18·22	29·57	63·45	2·37	4·11	5·03	6·28	8·22	12·32
0·7	6·00	13·15	19·22	26·56	43·10	92·49	3·42	5·92	7·25	9·04	11·83	17·75
0·8	9·74	21·33	29·57	43·10	69·93	150·1	5·48	9·49	11·62	14·49	18·97	28·46
0·9	20·89	45·77	63·45	92·49	150·1	322·0	11·62	20·13	24·65	30·74	40·25	60·37

Coefficient of variation

	(iii) $\rho = 2·5$, q_k						(iv) $\rho = 3·5$, q_k					
q_a	0·25	0·5	0·6	0·7	0·8	0·9	0·25	0·5	0·6	0·7	0·8	0·9
0·25	3·98	2·99	2·80	2·65	2·54	2·44	2·31	1·69	1·57	1·48	1·40	1·34
0·5	—	2·26	2·11	2·01	1·92	1·85	—	1·40	1·30	1·22	1·16	1·10
0·6	—	—	1·97	1·88	1·79	1·73	—	—	1·21	1·14	1·08	1·03
0·7	—	—	—	1·78	1·70	1·64	—	—	—	1·07	1·01	0·96
0·8	—	—	—	—	1·63	1·57	—	—	—	—	0·96	0·92
0·9	—	—	—	—	—	1·56	—	—	—	—	—	0·88

	(v) $\rho = 4·5$, q_k, k						(vi) Negative binomial $\{1/p_a - q_a A/p_a\}^{-k}$, $k(\rho \to \infty, a \to \infty)$					
	0·25	0·5	0·6	0·7	0·8	0·9						
q_a	1·5	4·5	6·75	10·5	18·0	40·5	1·5	4·5	6·75	10·5	18·0	40·5
0·25	2·13	1·53	1·43	1·33	1·26	1·20	1·63	0·94	0·77	0·62	0·47	0·31
0·5	1·53	1·12	1·04	0·97	0·92	0·88	1·15	0·67	0·54	0·44	0·33	0·22
0·6	1·43	1·04	0·96	0·90	0·85	0·81	1·05	0·61	0·50	0·40	0·30	0·20
0·7	1·33	0·97	0·90	0·84	0·80	0·76	0·98	0·56	0·48	0·37	0·28	0·19
0·8	1·26	0·92	0·85	0·80	0·76	0·72	0·91	0·53	0·43	0·35	0·26	0·18
0·9	1·20	0·88	0·81	0·76	0·72	0·69	0·86	0·50	0·41	0·33	0·25	0·17

The values of the coefficient of variation (c. of v.) increase rather more rapidly (by 60–80 per cent) than those of the S.D. as ρ falls from 3·5 to 2·5. This, of course, is due to the decrease of the mean. The values of the c. of v. for $\rho = 4{\cdot}5$ are always *less* (10–20 per cent less) than those for $\rho = 3{\cdot}5$.

The negative binomials are much less variable both absolutely and relatively than the corresponding G.W. distributions with the same values of k and q_a. This becomes much more marked as q_a increases.

The results might suggest that the standard deviation is not an appropriate measure of scale and that neither the S.D. nor the c. of v. are very appropriate measures of variation for $\rho \leqslant 4$. Naturally they change very rapidly with ρ, since for $\rho \leqslant 2$ the S.D. becomes infinite; while the c. of v. is infinite for $1 < \rho \leqslant 2$ and indeterminate for $\rho \leqslant 1$.

In the applications which have so far been made of these long-tailed discrete frequency distributions, the scale was provided by the nature of the problem. If, for example, one is dealing with the distribution of filarial worms per mite, or of species per genus, the variable *must* take the values $(0, 1, 2, 3, \ldots)$ or $(1, 2, 3, \ldots)$ and cannot take non-integral values. Thus the "unit of measurement" is one worm or one species.

However, one might consider some phenomenon, such as accident occurrence, in which the only variable in the external conditions was the period of exposure to risk. In that case the number of occurrences would be proportional to the exposure period. It would be natural to take the mean (or perhaps the median) as giving a unit of measurement.

As a matter of fact, this particular difficulty is not likely to occur with accident distributions for, as far as experience goes, they have finite fourth moments and $\rho > 4$. But there might be phenomena which did give smaller values of ρ. In such cases, it might be better to take the median or perhaps the interquartile distance as a scale parameter.

6.5. These considerations suggest some remarks on the subject of "skewness". The measures of skewness most usually used are $\gamma_1 = \sqrt{\beta_1}$ and K. Pearson's measure (mean − mode)/S.D., which for his system of frequency curves is equal to $\sqrt{\beta_1}(\beta_2+3)/(5\beta_2 - 6\beta_1 - 9)$. This reduces to $\sqrt{\beta_1}$ for his Type III $(2\beta_2 - 3\beta_1 - 6 = 0)$, which is the Gamma distribution, the normal $(\beta_1 = 0, \beta_2 = 3)$ and the exponential $(\beta_1 = 4, \beta_2 = 9)$, which are particular cases of it.

Neither of these two measures can be regarded as appropriate for these long-tailed distributions in which any of the first four moments might be infinite. Indeed it seems difficult to quantify the idea of skewness in such cases. It seems better to be satisfied with "length of tail". This can be measured by any suitably chosen percentage point; or better still by the *ratio of any two suitably chosen percentage points.* If the latter choice be made, it will be unimportant which two percentage points are selected for the purpose. For in distributions such as these the tail frequencies decline approximately in geometric progression. Whatever the choice, the results for the distributions compared will, for all choices, bear approximately the same proportion to one another. This can be done in either of the two cases mentioned at the end of Section 6.4. The unit of measurement is either unity or some suitably chosen scale parameter.

6.6. Although the standard deviation seems an undesirable measure of scale because it changes so rapidly for small values of ρ, it has one unexpected merit for $\rho \geqslant 2{\cdot}5$ when ρ is kept fixed. The percentage points, in terms of the standard deviation of the distributions for fixed ρ and varying q_a, q_k, are almost independent of the latter two parameters. This point is considered again below when the table of percentage points is discussed.

6 (CONT.). THE UPPER PERCENTAGE POINTS

6.71. The percentage points are given in Table 3. If $F(x)$ is the cumulative probability from 0 to x inclusive, the β per cent point is taken as $x\,(x \neq 0)$ if $1-F(x-1) > \beta/100 \geqslant 1-F(x)$. It is taken as 0 if $\beta/100 > 1-F(0)$; this happens in the tables in only one case—when $\rho = 0{\cdot}5, q_a = q_k = 0{\cdot}25$.

Not more than 2,000 terms were worked out on the electronic computer. However, the ratios of the successive percentage points exhibit a remarkable degree of constancy for fixed ρ and all values of q_a, q_k. Thus, it was possible to extrapolate (using a desk machine) in order to obtain the missing percentage points approximately; in all cases for $\rho = 1{\cdot}5$, 2·5, 3·5 and 4·5; but, for $\rho = 0{\cdot}5$, only the 10, 5, 2·5 and 1 per cent points could be obtained. Great accuracy is not important, but it is desired to indicate the order of magnitude of these values. They are shown in square brackets in Table 3.

(i) For $\rho = 2{\cdot}5$, 3·5 and 4·5 the extrapolated values are believed to be correct within ± 50 (except for the 0·005 per cent point at $\rho = 2{\cdot}5, q_a = q_k = 0{\cdot}9$ which is within ± 100). For $\rho = 1{\cdot}5$ the extrapolated 1 per cent point (for $q_a = q_k = 0{\cdot}9$) is believed to be within 0·5 per cent; the extrapolated 0·1 and 0·01 per cent points within 2 per cent (within 1 per cent for q_a, $q_k = 0{\cdot}25$ and 0·5) and the extrapolated 0·05 per cent points within 5 per cent (again within 1 per cent for $q_a, q_k = 0{\cdot}25$ and 0·5).

(ii) For $\rho = 0{\cdot}5$, the extrapolated 10 per cent point for $q_a = q_k = 0{\cdot}9$ is believed to be correct within ± 50 or 2·2 per cent; the extrapolated 5 per cent points within ± 50 for $q_a = 0{\cdot}7$, 0·8 with $q_k = 0{\cdot}9$, within 3 per cent for $q_a = 0{\cdot}9, q_k = 0{\cdot}9$. The extrapolated $2\frac{1}{2}$ per cent points are believed to be correct within ± 50 for $q_a = 0{\cdot}5$ and 0·6 with $q_k = 0{\cdot}9$ as well as for $q_a = 0{\cdot}7$, $q_k = 0{\cdot}8$ and 0·9, within 3 per cent for $q_a = 0{\cdot}8$, $q_k = 0{\cdot}9$ and within 5 per cent for $q_a = 0{\cdot}9, q_k = 0{\cdot}9$. The extrapolated 1 per cent points are believed to be correct within ± 50 for $q_a = 0{\cdot}25, q_k = 0{\cdot}9$, for $q_a = 0{\cdot}5, q_k = 0{\cdot}6$, 0·7, 0·8 and within 2 per cent for $q_a = 0{\cdot}5, q_k = 0{\cdot}9$.

For q_a, $q_k = 0{\cdot}6$, 0·7, 0·8, 0·9, the 1 per cent point has been taken to be 6·5 times the $2\frac{1}{2}$ per cent point, rounded off to the nearest 50, for values less than 10,000; and to three significant figures for five- and six-figure values. The ratio is certainly between 6 and 7 in all these cases. This makes the error $\leqslant 8$ per cent, except $q_a = 0{\cdot}8$, $q_k = 0{\cdot}9$ ($\leqslant 10{\cdot}2$ per cent) and $q_a = 0{\cdot}9, q_k = 0{\cdot}9$ ($\leqslant 12{\cdot}1$ per cent). For $\rho = 2{\cdot}5$, 3·5 and 4·5, the percentage points are given both in "actual measure" ($x = 0, 1, 2, 3$, etc.) and in "standard measure" $\{X = (x-\mu)/\sigma\}$. For $\rho \leqslant 2$ the S.D. is infinite and for $\rho \leqslant 1$, the mean is infinite. For $\rho = 0{\cdot}5$ and 1·5, therefore, the percentage points can only be given in "actual measure".

The G.W. distributions for $\rho = 0{\cdot}25$, 0·5, (1) 4·5, are considered in Sections 6.72, 6.73; the negative binomials in 6.74, 6.75.

6.72. We now consider the percentage points in "actual measure". The following features are of interest:

(a) The extreme length of tail when $\rho = 0{\cdot}5$ and 1·5 and the great rate at which this increases as ρ drops from 1·5 to 0·5 are especially striking. The second phenomenon becomes increasingly noticeable as the percentage β corresponding to the percentage point x_β diminishes, q_a, q_k being fixed; and for a fixed percentage point as q_a, q_k increase. For example, when $\rho = 0{\cdot}5, q_a = 0{\cdot}25, q_k = 0{\cdot}9$, even the 1 per cent point is 3,200; but it is only 156, 108, 99, 98 for $\rho = 1{\cdot}5$, $2{\cdot}5$, $3{\cdot}5$, $4{\cdot}5$ respectively. When $\rho = 0{\cdot}5, q_a = 0{\cdot}9, q_k = 0{\cdot}9$, the 1 per cent point is 238,000 compared with 3,300 when $\rho = 1{\cdot}5, q_a = 0{\cdot}9, q_k = 0{\cdot}9$.

(b) Quite generally, for fixed ρ, all percentage points increase rapidly with q_a for fixed q_k and vice versa. The ratios (x_β for $q_k = 0{\cdot}9$)/(x_β for $q_k = 0{\cdot}25$) for fixed q_a, all tabulated β and fixed ρ are almost constant. When $\rho = 0{\cdot}5$, $1{\cdot}5$, $2{\cdot}5$, $3{\cdot}5$, $4{\cdot}5$ they average respectively 74, 22, 18, 17, 17. The highest and lowest values of these ratios are respectively 84, 27, 29, 22, 21 and 71, 19, 16, 15, 15.

The value 84 is for $q_a = 0{\cdot}5$. For $q_a = q_k = 0{\cdot}25$, the 10 per cent point is zero, which gives an infinite ratio. The highest ratios all occur at a 10 per cent point, and only occur once. The lowest always occur at a 0·005 per cent point and may occur several times. When $\rho = 4{\cdot}5$ the ratio is 15 at all 0·01 and 0·005 per cent points. The same is true for $\rho = 3{\cdot}5$ and $2{\cdot}5$, with one exception in each case, when it is 16.

(c) The way in which a fixed percentage point changes with ρ, when q_a, q_k are fixed is of considerable interest.

As $\rho \to \infty$ the distribution is asymptotically normal with a very large mean and S.D., so that, in "actual measure", any upper percentage point $x_\beta \to \infty$ for all β. When $\rho = 0$, $F(0) = 1$, all the frequency is at $x = 0$, and all the percentage points, as here defined, are zero. If however $\rho = \varepsilon$ (say), where ε is very small, it is not difficult to see that the frequency at $x = 0$ is very nearly unity, and ~ 1 as $\varepsilon \to 0$. Any other frequency $f_r \sim \gamma\varepsilon/r$ as $\varepsilon \to 0$, where $\gamma = q_a q_k/(1-q_a q_k)$. As $\rho \to 0$, in fact, the distribution tends asymptotically to coincide with the axes, from the point (0, 1) vertically down to the oirigin and along the axis of x.

Thus, any upper percentage point is zero if $\beta/100 > (1-F(0))$. If $\rho = \rho_0(\beta)$ when $\beta/100 = 1-F(0)$, x_β will increase with ρ, at least for a time, for $\rho > \rho_0$. On the other hand, for any sufficiently small $\beta < 100\,(1-F(0))$, x_β will increase as ρ *diminishes* to some value $\rho_1(\beta)$.

Thus, any upper percentage point x_β will be zero as long as $\rho \leqslant \rho_0(\beta)$ and will tend to infinity with ρ. If β is sufficiently small, x_β will increase with ρ from $\rho = \rho_0(\beta)$ to a maximum at some value of $\rho (= \rho_1(\beta)$, say) then decline to a minimum and finally rise again, tending to infinity with ρ. On the other hand, if β is sufficiently large, x_β will increase monotonically as ρ increases from $\rho_0(\beta)$ to infinity. This is certainly true for the median, which is given by $I\{(\rho\alpha-1)(\rho\kappa-1)/(\rho+1)\}$ where $\alpha = q_a/p_a$, $\kappa = q_k/p_k$.

It is conjectured therefore that there is some definite value of β above which x_β increases monotonically with ρ and below which x_β has a maximum (at $\rho_1(\beta)$, say) and a minimum (at $\rho_2(\beta)$, say) (where $\rho_1 > \rho_2$). This critical value of β as well as $\rho_0(\beta)$, $\rho_1(\beta)$, $\rho_2(\beta)$, depend also on q_a, q_k.

Table 3 is consistent with this conclusion. In the following discussion q_a and q_k may, of course, be interchanged.

(a) The 10 per cent points *rise* as ρ increases from 0·5 to 4·5 for $q_a = 0{\cdot}25, q_k = 0{\cdot}25$ and 0·5. For $q_a = 0{\cdot}25$, $q_k = 0{\cdot}6$ and 0·7, the values are respectively 4, 4, 5, 6, 7, 7, 7, 8, 9, 11. For $q_a = 0{\cdot}25$, $q_k = 0{\cdot}8$, 0·9 and for all other pairs of values they fall to a minimum close to $\rho = 1{\cdot}5$ and then rise again. Since all percentage points are zero for sufficiently small ρ, there must be a maximum for $\rho \leqslant 0{\cdot}5$ in all these cases (except when $\rho = 0{\cdot}5$, $q_a = q_k = 0{\cdot}25$, where the 10 per cent point is zero).

TABLE 3

Absolute and relative upper percentage points for long-tailed G.W. distribution
[x = *Absolute* % *point*, $X = \{(x-\mu)/\sigma\}$]

			$\rho = 0{\cdot}5$					
			q_k					
q_a			0·25	0·5	0·6	0·7	0·8	0·9
	10%	x	0	2	4	7	12	30
	5%	x	2	9	16	28	52	126
	2·5%	x	7	38	66	114	212	511
0·25	1%	x	45	240	412	713	1,331	[3,200]
	0·1%	x						
	0·01%	x						
	0·005%	x						
	10%	x	2	12	22	38	70	169
	5%	x	9	51	88	153	285	687
	2·5%	x	38	206	354	613	1,144	[2,750]
0·5	1%	x	240	1,291	[2,250]	[3,900]	[7,300]	[17,500]
	0·1%	x						
	0·01%	x						
	0·005%	x						
	10%	x	4	22	37	65	121	293
	5%	x	16	88	152	263	491	1,182
	2·5%	x	66	354	609	1,054	1,968	[4,750]
0·6	1%	x	412	[2,250]	[3,950]	[6,850]	[12,800]	[30,800]
	0·1%	x						
	0·01%	x						
	0·005%	x						
	10%	x	7	38	65	113	211	508
	5%	x	28	153	263	455	850	[2,050]
	2·5%	x	114	613	1,054	1,825	[3,400]	[8,150]
0·7	1%	x	713	[3,900]	[6,850]	[11,900]	[22,200]	[52,800]
	0·1%	x						
	0·01%	x						
	0·005%	x						
	10%	x	12	70	121	211	394	950
	5%	x	52	285	491	850	1,587	[3,900]
	2·5%	x	212	1,144	1,968	[3,400]	[6,350]	[15,200]
0·8	1%	x	1,331	[7,300]	[12,800]	[22,200]	[41,300]	[99,100]
	0·1%	x						
	0·01%	x						
	0·005%	x						
	10%	x	30	169	293	508	950	[2,300]
	5%	x	126	687	1,182	[2,050]	[3,900]	[9,150]
	2·5%	x	511	[2,750]	[4,750]	[8,100]	[15,200]	[36,500]
0·9	1%	x	[3,200]	[11,500]	[30,800]	[52,800]	[99,100]	[238,000]
	0·1%	x						
	0·01%	x						
	0·005%	x						

TABLE 3 (*cont.*)

			$\rho = 1{\cdot}5$					
			q_k					
q_a			0·25	0·5	0·6	0·7	0·8	0·9
	10%	x	1	3	4	7	11	25
	5%	x	2	6	8	12	21	46
	2·5%	x	4	10	14	21	36	80
0·25	1%	x	7	19	25	42	71	156
	0·1%	x	38	96	138	208	347	765
	0·01%	x	198	450	650	977	1,630	[3,600]
	0·005%	x	284	716	1,037	1,555	[2,600]	[5,700]
	10%	x	3	9	13	19	33	73
	5%	x	6	15	22	34	57	127
	2·5%	x	10	26	38	57	96	211
0·5	1%	x	19	50	72	109	183	404
	0·1%	x	96	242	349	526	880	1,939
	0·01%	x	450	1,135	1,640	[2,450]	[4,150]	[9,100]
	0·005%	x	716	1,805	[2,600]	[3,900]	[6,550]	[14,400]
	10%	x	4	13	19	29	49	109
	5%	x	8	22	33	50	84	186
	2·5%	x	14	38	55	83	139	308
0·6	1%	x	25	72	105	158	265	585
	0·1%	x	138	349	504	760	1,270	[2,800]
	0·01%	x	650	1,640	[2,300]	[3,500]	[5,800]	[13,000]
	0·005%	x	1,037	[2,600]	[3,500]	[5,400]	[9,100]	[20,000]
	10%	x	7	19	29	44	75	167
	5%	x	12	34	50	76	128	283
	2·5%	x	21	57	83	126	211	461
0·7	1%	x	42	109	158	239	401	886
	0·1%	x	208	526	760	1,145	1,915	[4,250]
	0·1%	x	977	[2,450]	[3,500]	[5,300]	[8,800]	[19,500]
	0·005%	x	1,555	[3,900]	[5,400]	[8,200]	[13,700]	[30,000]
	10%	x	11	33	49	75	127	284
	5%	x	21	57	84	128	215	478
	2·5%	x	36	96	139	211	355	785
0·8	1%	x	71	183	265	401	672	1,484
	0·1%	x	347	880	1,270	1,915	[3,200]	[7,100]
	0·1%	x	1,630	[4,150]	[5,800]	[8,800]	[14,600]	[32,000]
	0·005%	x	[2,600]	[6,550]	[9,100]	[13,700]	[23,000]	[50,000]
	10%	x	25	73	109	167	284	633
	5%	x	46	127	186	283	478	1,060
	2·5%	x	80	211	308	467	785	1,738
0·9	1%	x	156	404	585	886	1,484	[6,300]
	0·1%	x	765	1,939	[2,800]	[4,250]	[7,100]	[15,700]
	0·01%	x	[3,600]	[9,100]	[13,000]	[19,500]	[32,000]	[72,000]
	0·005%	x	[5,700]	[14,400]	[20,000]	[30,000]	[50,000]	[111,000]

TABLE 3 (*cont.*)

				$\rho = 2{\cdot}5$											
				q_k											
q_a				0·25		0·5		0·6		0·7		0·8		0·9	
	10%	x	X	1	0·3	4	0·6	5	0·5	8	0·6	13	0·5	29	0·5
	5%	x	X	2	0·8	6	1·1	8	1·0	13	1·1	21	1·1	46	1·1
	2·5%	x	X	4	1·9	9	1·8	13	1·9	19	1·8	31	1·8	68	1·8
0·25	1%	x	X	6	3·0	14	3·0	20	3·1	30	3·1	49	3·1	108	3·1
	0·1%	x	X	18	9·6	41	9·5	58	9·6	85	9·5	140	9·5	304	9·5
	0·01%	x	X	49	26·5	109	25·9	152	25·8	225	25·8	369	25·7	799	25·7
	0·005%	x	X	65	35·3	145	34·5	202	34·4	299	34·4	490	34·4	1,062	34·4
	10%	x	X	4	0·6	10	0·6	14	0·6	21	0·6	36	0·6	79	0·6
	5%	x	X	6	1·1	15	1·2	21	1·1	32	1·1	53	1·1	116	1·1
	2·5%	x	X	9	1·8	21	1·8	31	1·9	46	1·9	75	1·8	165	1·8
0·5	1%	x	X	14	3·0	34	3·2	48	3·2	71	3·1	116	3·1	254	3·1
	0·1%	x	X	41	9·5	93	9·4	131	9·4	194	9·4	318	9·4	690	9·4
	0·01%	x	X	109	25·9	244	25·4	341	25·3	503	25·3	824	25·2	1,787	25·2
	0·005%	x	X	145	34·5	323	33·8	452	33·7	667	33·7	1,093	33·7	[2,350	33·6]
	10%	x	X	5	0·5	14	0·6	20	0·6	31	0·6	52	0·6	115	0·6
	5%	x	X	8	1·0	21	1·1	30	1·1	46	1·1	76	1·1	117	1·1
	2·5%	x	X	13	1·9	31	1·9	43	1·8	65	1·8	108	1·8	235	1·8
0·6	1%	x	X	20	3·1	48	3·2	67	3·1	100	3·1	165	3·1	360	3·1
	0·1%	x	X	58	9·6	131	9·4	184	9·4	272	9·4	447	9·4	970	9·4
	0·01%	x	X	152	25·7	341	25·3	477	25·2	704	25·2	1,154	25·2	[2,500	25·1]
	0·005%	x	X	202	34·3	452	33·7	632	33·6	933	33·6	1,529	33·5	[3,300	33·5]
	10%	x	X	8	0·6	21	0·6	31	0·6	47	0·6	79	0·6	175	0·6
	5%	x	X	13	1·1	32	1·1	46	1·1	69	1·1	114	1·1	251	1·1
	2·5%	x	X	19	1·8	46	1·9	65	1·8	97	1·8	161	1·8	352	1·8
0·7	1%	x	X	30	3·1	71	3·1	100	3·1	149	3·1	245	3·1	535	3·1
	0·1%	x	X	85	9·5	194	9·4	272	9·4	402	9·4	660	9·4	1,434	9·4
	0·01%	x	X	225	25·8	503	25·3	704	25·2	1,038	25·2	1,703	25·2	[3,700	25·1]
	0·005%	x	X	299	34·4	667	33·7	933	33·6	1,377	33·6	[2,250	33·5]	[4,850	33·4]
	10%	x	X	13	0·5	36	0·6	52	0·6	79	0·6	133	0·6	293	0·6
	5%	x	X	21	1·1	53	1·1	76	1·1	114	1·1	191	1·1	419	1·1
	2·5%	x	X	31	1·8	75	1·8	108	1·8	161	1·8	267	1·8	583	1·8
0·8	1%	x	X	49	3·1	116	3·1	165	3·1	245	3·1	405	3·1	883	3·1
	0·1%	x	X	140	9·5	319	9·4	447	9·4	660	9·4	1,085	9·4	[2,350	9·4]
	0·01%	x	X	368	25·7	824	25·2	1,154	25·2	1,703	25·2	[2,800	25·1]	[6,050	25·1]
	0·005%	x	X	490	34·4	1,093	33·7	1,529	33·5	[2,250	33·5]	[3,700	33·4]	[7,950	33·4]
	10%	x	X	29	0·5	79	0·6	115	0·6	175	0·6	293	0·6	648	0·6
	5%	x	X	46	1·1	116	1·1	167	1·1	251	1·1	419	1·1	920	1·1
	2·5%	x	X	68	1·8	165	1·8	235	1·8	352	1·8	583	1·8	1,277	1·8
0·9	1%	x	X	108	3·1	254	3·1	360	3·1	535	3·1	883	3·1	1,926	3·0
	0·1%	x	X	304	9·5	690	9·4	970	9·4	1,434	9·4	[2,350	9·4]	[5,150	9·4]
	0·01%	x	X	799	25·7	1,787	25·2	[2,500	25·1]	[3,700	25·1]	[6,050	25·1]	[13,150	25·1]
	0·005%	x	X	1,062	34·4	[2,350	33·6]	[3,300	33·5]	[4,850	33·4]	[7,950	33·4]	[17,200	33·4]

TABLE 3 (*cont.*)

$\rho = 3{\cdot}5$

q_a				q_b 0·25		0·5		0·6		0·7		0·8		0·9	
	10%	x	X	2	1·2	4	0·9	6	0·9	9	1·0	16	1·0	34	1·0
	5%	x	X	3	2·0	6	1·7	9	1·8	14	1·8	23	1·7	49	1·7
	2·5%	x	X	4	2·8	9	2·8	13	2·7	19	2·7	31	2·7	68	2·7
0·25	1%	x	X	6	4·4	13	4·3	19	4·3	28	4·3	46	4·3	99	4·3
	0·1%	x	X	14	10·8	31	10·6	43	10·5	63	10·5	103	10·6	223	10·6
	0·01%	x	X	30	23·6	65	22·8	90	22·8	132	22·8	215	22·7	463	22·7
	0·005%	x	X	38	30·0	81	28·5	112	28·3	163	28·2	265	28·3	572	28·3
	10%	x	X	4	0·9	11	0·9	16	0·9	24	0·9	41	0·9	89	0·9
	5%	x	X	6	1·7	16	1·6	23	1·6	34	1·6	56	1·6	122	1·6
	2·5%	x	X	9	2·8	21	2·4	30	2·4	45	2·4	74	2·4	162	2·4
0·5	1%	x	X	13	4·3	31	3·8	43	3·8	64	3·8	105	3·8	227	3·7
	0·1%	x	X	31	10·6	68	9·2	95	9·2	139	9·1	227	9·1	490	9·1
	0·01%	x	X	65	22·8	140	19·7	194	19·5	284	19·5	462	19·4	996	19·4
	0·005%	x	X	81	28·5	173	24·5	239	24·3	350	24·2	569	24·1	1,226	24·1
	10%	x	X	6	0·9	16	0·9	23	0·9	35	0·9	59	0·9	130	0·9
	5%	x	X	9	1·8	23	1·6	32	1·6	48	1·6	80	1·6	175	1·6
	2·5%	x	X	13	2·7	30	2·4	43	2·4	64	2·4	105	2·4	230	2·4
0·6	1%	x	X	19	4·3	43	3·8	61	3·8	90	3·8	147	3·7	320	3·7
	0·1%	x	X	43	10·5	95	9·2	132	9·1	194	9·1	316	9·1	684	9·1
	0·01%	x	X	90	22·8	194	19·5	269	19·4	394	19·4	642	19·3	1,384	19·3
	0·005%	x	X	112	28·3	239	24·3	332	24·2	485	24·1	790	24·0	1,702	24·0
	10%	x	X	9	1·0	24	0·9	35	0·9	53	0·9	89	0·9	196	0·9
	5%	x	X	14	1·8	34	1·6	48	1·6	72	1·6	120	1·6	262	1·6
	2·5%	x	X	19	2·7	45	2·4	64	2·4	95	2·4	157	2·4	342	2·4
0·7	1%	x	X	28	4·3	64	3·8	90	3·8	133	3·8	218	3·7	474	3·7
	0·1%	x	X	63	10·5	139	9·1	194	9·1	284	9·1	464	9·0	1,004	9·0
	0·01%	x	X	132	22·8	284	19·5	394	19·4	576	19·3	939	19·3	[2,050	19·3]
	0·005%	x	X	163	28·2	350	24·2	485	24·1	709	24·0	1,154	23·9	[2,500	23·8]
	10%	x	X	16	1·0	41	0·9	59	0·9	89	0·9	149	0·9	328	0·9
	5%	x	X	23	1·7	56	1·6	80	1·6	120	1·6	199	1·6	436	1·6
	2·5%	x	X	31	2·7	74	2·4	105	2·4	157	2·4	259	2·4	566	2·4
0·8	1%	x	X	46	4·3	105	3·8	147	3·7	218	3·7	359	3·7	781	3·7
	0·1%	x	X	103	10·6	227	9·1	316	9·1	464	9·0	759	9·0	1,641	9·0
	0·01%	x	X	215	22·7	462	19·4	642	19·3	939	19·3	1,530	19·2	[3,300	19·2]
	0·005%	x	X	265	28·3	569	24·1	790	24·0	1,154	23·9	1,880	23·9	[4,050	23·6]
	10%	x	X	34	1·0	89	0·9	130	0·9	196	0·9	328	0·9	724	0·9
	5%	x	X	49	1·7	122	1·6	175	1·6	262	1·6	436	1·6	957	1·6
	2·5%	x	X	68	2·7	162	2·4	230	2·4	342	2·4	566	2·4	1,236	2·4
0·9	1%	x	X	99	4·3	227	3·7	320	3·7	474	3·7	781	3·7	1,698	3·7
	0·1%	x	X	223	10·6	490	9·1	684	9·1	1,004	0·9	1,641	9·0	[3,550	9·0]
	0·01%	x	X	463	22·7	996	19·4	1,384	19·3	[2,050	19·3]	[3,300	19·2]	[7,100	19·2]
	0·005%	x	X	572	28·3	1,226	24·1	1,702	24·0	[2,500	23·8]	[4,050	23·6]	[8,700	23·4]

TABLE 3 (*cont.*)

				$\rho = 4{\cdot}5$											
				q_k, k											
				0·25		0·5		0·6		0·7		0·8		0·9	
				1·5		4·5		6·75		10·5		18·0		40·5	
q_a															
	10%	x	X	2	1·0	5	1·0	7	1·0	11	1·1	18	1·1	39	1·0
	5%	x	X	3	1·7	7	1·7	10	1·7	15	1·8	25	1·8	54	1·8
	2·5%	x	X	4	2·5	10	2·7	13	2·4	20	2·6	33	2·6	71	2·6
0·25	1%	x	X	6	3·9	13	3·7	19	3·9	28	3·9	45	3·8	98	3·9
	0·1%	x	X	13	9·1	27	8·4	38	8·5	56	8·6	91	8·6	195	8·5
	0·01%	x	X	24	17·2	51	16·5	70	16·3	102	16·2	166	16·2	357	16·3
	0·005%	x	X	29	20·8	61	19·9	84	19·7	122	19·6	198	19·5	424	19·5
	10%	x	X	5	1·0	13	1·1	18	1·0	27	1·0	46	1·1	101	1·1
	5%	x	X	7	1·7	17	1·7	24	1·7	36	1·7	60	1·7	132	1·8
	2·5%	x	X	10	2·7	22	2·5	32	2·6	47	2·6	77	2·5	168	2·5
0·5	1%	x	X	13	3·7	30	3·7	43	3·8	63	3·8	103	3·7	224	3·8
	0·1%	x	X	27	8·4	59	8·2	83	8·2	121	8·2	197	8·2	425	8·2
	0·01%	x	X	51	16·5	108	15·7	149	15·6	217	15·5	353	15·5	758	15·4
	0·005%	x	X	61	19·9	128	18·8	177	18·7	257	18·5	417	18·5	897	18·5
	10%	x	X	7	1·0	18	1·0	26	1·0	40	1·1	66	1·1	146	1·1
	5%	x	X	10	1·7	24	1·7	35	1·8	52	1·7	86	1·7	189	1·8
	2·5%	x	X	13	2·4	32	2·6	45	2·6	66	2·5	109	2·5	238	2·5
0·6	1%	x	X	19	3·9	43	3·8	60	3·8	88	3·7	145	3·7	315	3·7
	0·1%	x	X	38	8·5	83	8·2	115	8·2	168	8·1	274	8·1	591	8·1
	0·01%	x	X	70	16·3	149	15·6	206	15·4	300	15·4	488	15·3	1,049	15·3
	0·005%	x	X	84	19·7	177	18·7	244	18·5	355	18·4	577	18·3	1,240	18·3
	10%	x	X	11	1·1	27	1·0	40	1·1	60	1·1	100	1·1	221	1·1
	5%	x	X	15	1·8	36	1·7	52	1·7	78	1·8	129	1·7	283	1·8
	2·5%	x	X	20	2·6	47	2·6	66	2·5	99	2·5	163	2·5	355	2·5
0·7	1%	x	X	28	3·9	63	3·8	88	3·7	131	3·8	215	3·7	466	3·7
	0·1%	x	X	56	8·6	121	8·2	168	8·1	246	8·1	402	8·1	867	8·1
	0·01%	x	X	102	16·2	217	15·5	300	15·4	438	15·3	712	15·3	1,531	15·2
	0·005%	x	X	122	19·6	257	18·5	355	18·4	517	18·3	840	18·2	1,807	18·2
	10%	x	X	18	1·1	46	1·1	66	1·1	100	1·1	168	1·1	369	1·1
	5%	x	X	25	1·8	60	1·7	86	1·7	129	1·7	215	1·7	470	1·7
	2·5%	x	X	33	2·6	77	2·5	109	2·5	163	2·5	269	2·5	586	2·5
0·8	1%	x	X	45	3·8	103	3·7	145	3·7	215	3·7	353	3·7	767	3·7
	0·1%	x	X	91	8·6	197	8·2	274	8·1	402	8·1	655	8·0	1,415	8·0
	0·01%	x	X	166	16·2	353	15·5	488	15·3	712	15·3	1,157	15·2	[2,500	15·1]
	0·005%	x	X	198	19·5	417	18·5	577	18·3	840	18·2	1,365	18·2	[2,900	18·1]
	10%	x	X	39	1·0	101	1·1	146	1·1	221	1·1	369	1·1	814	1·1
	5%	x	X	54	1·8	132	1·8	189	1·8	283	1·8	470	1·7	1,031	1·8
	2·5%	x	X	71	2·6	168	2·5	238	2·5	355	2·5	586	2·5	1,279	2·5
0·9	1%	x	X	98	3·9	224	3·8	315	3·7	466	3·7	767	3·7	1,667	3·7
	0·1%	x	X	195	8·5	425	8·2	591	8·1	867	8·1	1,415	8·0	[3,050	8·0]
	0·01%	x	X	357	16·3	758	15·4	1,049	15·3	1,531	15·2	[2,500	15·1]	[5,350	15·0]
	0·005%	x	X	424	19·5	897	18·5	1,240	18·3	1,807	18·2	[2,900	18·1]	[6,300	18·0]

TABLE 3 (*cont.*)

				Negative binomial $\{1/p_a-(q_a\,A/p_a)\}^{-k}$											
				k											
q_a				1·5		4·5		6·75		10·5		18·0		40·5	
	10%	*x*	*X*	2	1·8	3	1·1	5	1·6	6	1·2	10	1·4	19	1·3
	5%	*x*	*X*	2	1·8	4	1·8	5	1·6	7	1·6	11	1·8	21	1·8
	2·5%	*x*	*X*	3	3·1	5	2·5	6	2·2	8	2·1	12	2·1	23	2·2
0·25	1%	*x*	*X*	3	3·1	6	3·2	7	2·7	10	3·1	14	2·8	25	2·7
	0·1%	*x*	*X*	5	5·5	8	4·6	10	4·5	12	3·9	17	3·9	29	3·7
	0·01%	*x*	*X*	7	8·0	10	6·0	12	5·6	15	5·3	20	5·0	33	4·6
	0·005%	*x*	*X*	7	8·0	11	6·7	13	6·2	16	5·8	21	5·3	34	4·8
	10%	*x*	*X*	4	1·4	9	1·5	12	1·4	17	1·4	26	1·3	52	1·3
	5%	*x*	*X*	5	2·0	10	1·8	14	2·0	19	1·9	29	1·8	56	1·7
	2·5%	*x*	*X*	6	2·6	12	2·5	15	2·2	21	2·3	31	2·2	60	2·2
0·5	1%	*x*	*X*	7	3·2	14	3·2	17	2·8	23	2·7	34	2·7	64	2·6
	0·1%	*x*	*X*	11	5·5	18	4·5	22	4·2	29	4·0	41	3·8	73	3·6
	0·01%	*x*	*X*	14	7·2	22	5·8	27	5·5	34	5·1	47	4·8	80	4·4
	0·005%	*x*	*X*	16	8·4	23	6·2	28	5·8	36	5·6	49	5·2	83	4·7
	10%	*x*	*X*	5	1·2	12	1·3	17	1·4	24	1·3	38	1·3	77	1·3
	5%	*x*	*X*	7	2·0	14	1·8	19	1·8	27	1·8	42	1·8	82	1·7
	2·5%	*x*	*X*	8	2·4	16	2·3	22	2·4	30	2·3	45	2·2	87	2·1
0·6	1%	*x*	*X*	10	3·3	19	3·0	25	3·0	33	2·7	49	2·7	92	2·5
	0·1%	*x*	*X*	15	5·4	25	4·4	31	4·1	41	4·0	58	3·8	105	3·6
	0·01%	*x*	*X*	20	7·5	31	5·9	38	5·5	48	5·1	66	4·7	115	4·4
	0·005%	*x*	*X*	21	7·9	33	6·4	39	5·7	50	5·5	69	5·1	118	4·6
	10%	*x*	*X*	8	1·3	18	1·3	25	1·3	37	1·4	58	1·4	118	1·3
	5%	*x*	*X*	10	1·9	22	1·9	29	1·9	41	1·8	63	1·8	125	1·7
	2·5%	*x*	*X*	12	2·5	24	2·3	32	2·3	45	2·3	68	2·2	132	2·1
0·7	1%	*x*	*X*	15	3·4	28	3·0	37	3·0	50	2·8	74	2·7	140	2·6
	0·1%	*x*	*X*	22	5·4	37	4·5	46	4·2	61	4·0	87	3·8	157	3·7
	0·01%	*x*	*X*	29	7·5	45	5·8	55	5·4	71	5·1	98	4·7	173	4·5
	0·005%	*x*	*X*	31	8·1	47	6·2	58	5·8	73	5·5	102	5·1	177	4·6
	10%	*x*	*X*	13	1·3	31	1·4	43	1·4	61	1·3	97	1·3	199	1·3
	5%	*x*	*X*	17	2·0	36	1·9	48	1·8	68	1·8	106	1·8	211	1·7
	2·5%	*x*	*X*	20	2·6	40	2·3	54	2·3	74	2·2	113	2·2	222	2·1
0·8	1%	*x*	*X*	25	3·5	46	3·0	60	2·8	82	2·8	123	2·7	235	2·6
	0·1%	*x*	*X*	36	5·5	60	4·4	76	4·2	101	4·1	144	3·8	263	3·5
	0·1%	*x*	*X*	47	7·5	73	5·8	90	5·4	116	5·1	162	4·7	287	4·4
	0·005%	*x*	*X*	50	8·0	77	6·2	94	5·8	120	5·4	168	5·1	294	4·6
	10%	*x*	*X*	29	1·3	67	1·3	94	1·3	135	1·3	215	1·3	444	1·3
	5%	*x*	*X*	36	1·9	78	1·9	106	1·8	150	1·8	233	1·8	469	1·7
	2·5%	*x*	*X*	44	2·5	88	2·4	117	2·3	163	2·2	249	2·2	492	2·1
0·9	1%	*x*	*X*	53	3·2	101	3·0	132	2·9	180	2·8	269	2·7	519	2·6
	0·1%	*x*	*X*	76	5·2	130	4·4	164	4·2	217	4·0	314	3·8	578	3·5
	0·01%	*x*	*X*	99	7·2	158	5·8	195	5·4	251	5·1	354	4·8	631	4·4
	0·005%	*x*	*X*	106	8·0	166	6·2	204	5·8	261	5·4	365	5·0	645	4·6

(b) For all tabulated values of q_a, q_k the 5 and $2\frac{1}{2}$ per cent points decline, reach a minimum and then rise again as ρ increases from 0·5 to 4·5. The minimum occurs close to $\rho = 2{\cdot}5$ for the 5 per cent points and close to $\rho = 3{\cdot}5$ for the $2\frac{1}{2}$ per cent points. The 1 per cent points decline continuously between $\rho = 0{\cdot}5$ and $\rho = 4{\cdot}5$; but the values for $\rho = 3{\cdot}5$ and $\rho = 4{\cdot}5$ are very close together, particularly for small q_a or q_k. This suggests that the minimum occurs for a value of ρ not much greater than 4·5. The 0·1, 0·01 and 0·005 per cent points all decline continuously between $\rho = 0{\cdot}5$ and $\rho = 4{\cdot}5$. Since all the percentage points in this group *are not* zero for $\rho = 0{\cdot}5$, but *are* zero at $\rho = 0$ and $\to \infty$ as $\rho \to \infty$, there must be a maximum and a minimum for finite $\rho > 0$ in all these cases, but they are outside the range of the tables.

6.73. The most remarkable feature of the percentage points, when expressed in "standard measure" is the approximate constancy of their values for fixed ρ and all q_a, q_k. Even for fixed q_a, q_k and varying ρ, the changes in the three values are not large relative to their general level.

Table 4 may usefully be consulted in relation to the following discussion.

(a) Since the S.D. becomes infinite when $\rho = 2$, but the mean is finite down to $\rho = 1$, all standardized percentage points must be zero at $\rho = 2$. As $\rho \to \infty$, for given q_a, q_k they tend to their normal values. Thus any standardized percentage point must therefore rise to a maximum as ρ increases from $\rho = 2$. If this maximum is above the "normal" value, there need be no other maxima or minima for finite ρ. Table 3 suggests that there are not.

The 10 per cent points are all below their normal values. The 5 and $2\frac{1}{2}$ per cent points are below their normal values for $\rho = 2{\cdot}5$ and above for $\rho = 3{\cdot}5$ and 4·5. The 5 per cent points for $\rho = 4{\cdot}5$ are very close to those for 3·5, the $2\frac{1}{2}$ per cent points are smaller at $\rho = 4{\cdot}5$ than at $\rho = 3{\cdot}5$ for $q_a < 0{\cdot}5, q_k < 0{\cdot}5$ and a little larger in other cases; maxima near $\rho = 4$ are indicated.

The 1 per cent points are larger than their normal values; for all values of q_a, q_k, a maximum is indicated near $\rho = 3{\cdot}5$. The 0·1 per cent points are $2\frac{1}{2}$ or three times their normal values and also indicate a maximum near $\rho = 3{\cdot}5$. The 0·01 and 0·005 per cent points decrease with ρ up to $\rho = 4{\cdot}5$; but even at $\rho = 4{\cdot}5$ they are more than four times their normal values (about five times for the 0·005 per cent point). At $\rho = 2{\cdot}5$ the 0·01 per cent points are about seven times and the 0·005 per cent points about nine times their normal values. This indicates clearly the great length of tail of these distributions.

(b) The small changes of the percentage points for fixed ρ and varying q_a, q_k are also of some interest. In tabulating the percentage points, no attempt was made to "correct" for the discrete nature of the distributions. This means that, even at $\sigma = 40$, errors of 0·1 may occur in the differences between consecutive standardized values; that is to say, up to $(q_a, q_k) = (0{\cdot}5, 0{\cdot}8)$. When $q_a = 0{\cdot}25$, $q_k = 0{\cdot}25$ and 0·5, the error of the difference may be as much as 0·2.

Taking this into account, it seems that the small rise in the values of the 10 and 5 per cent points (from 0·3 to 0·6 and 0·8 to 1·1 respectively) for $\rho = 2{\cdot}5$, $q_a = 0{\cdot}25$, $q_k = 0{\cdot}25$ to 0·5 is genuine, but otherwise the values for constant ρ and increasing q_a, q_k, remain the same or decline slightly. There is a very slight decline in the values of the $2\frac{1}{2}$ and 1 per cent points. This decline is much more marked in the 0·1, 0·01 and 0·005 per cent points.

6.741. The negative binomials in the tables have been taken to be $\{1/p_a - (q_a A/p_a)\}^{-k}$ where k is finite and $k = 4{\cdot}5 q_k/p_k$, q_k taking the values 0·25, 0·5 (0·1) 0·9, i.e. those

TABLE 4

Upper percentage points in standard measure for low and high values of q_a, q_k

Percentage point	Normal values		$\rho = 2{\cdot}5$				$\rho = 3{\cdot}5$				$\rho = 4{\cdot}5$			
		q_a	0·25	0·25	0·25	0·9	0·25	0·25	0·25	0·9	0·25	0·25	0·25	0·9
		q_k	0·25	0·5	0·9	0·9	0·25	0·5	0·9	0·9	0·25	0·5	0·9	0·9
10	1·3		0·3	0·6	0·5	0·6	1·2	0·9	1·0	0·9	1·0	1·0	1·0	1·1
5	1·6		0·8	1·1	1·1	1·1	2·0	1·7	1·7	1·6	1·7	1·7	1·8	1·8
2·5	2·0		1·9	1·8	1·8	1·8	2·8	2·8	2·7	2·4	2·5	2·7	2·6	2·5
1	2·3		3·0	3·0	3·1	3·0	4·4	4·3	4·3	3·7	3·9	3·7	3·9	3·7
0·1	3·1		9·6	9·5	9·5	9·4	10·8	10·6	10·6	9·0	9·1	8·4	8·5	8·0
0·01	3·7		26·5	25·9	25·7	25·1	23·6	22·8	22·7	19·2	17·2	16·5	16·3	15·0
0·005	3·9		35·3	34·5	34·4	33·4	30·0	28·5	28·3	23·4	20·8	19·9	19·5	18·0

appropriate for the G.W. distributions with $\rho = 4{\cdot}5$. In the negative binomials themselves $q_k = k/(k+\rho) = 0$ since $\rho \to \infty$ and k is finite. This has been explained before but is repeated here for clarity. We have also noted already that the negative binomials obtained by interchanging the arguments q_k, q_a are *not* the same, except in the mean, unless $q_a = q_k$, when they are identical.

Thus the tabulated negative binomials can be divided into two classes, (i) those for which $q_k \geqslant q_a$, and (ii) those for which $q_a \geqslant q_k$. Those in which $q_a = q_k$, are here included in both classes. The percentage points in (ii) are greater than the corresponding ones in (i) except when $q_a = q_k$, when they are of course equal. However, the difference between them is usually small; for fixed q_a in (i) it increases as q_k increases; for fixed q_k in (i) it decreases as q_a increases. The position is best summarized by examining the *ratios* of corresponding values in (ii) and (i). For each percentage point there are 21 pairs of values for the arguments $q_a, q_k = 0{\cdot}25$, $0{\cdot}5\,(0{\cdot}1)\,0{\cdot}9$. The ratio of the 21 corresponding values in (ii) and (i), ((ii)/(i)) has a J-shaped distribution with the greatest frequency at 1·0. The greatest values and median are as follows:

Percentage point	*Greatest value of ratio*	*Median ratio*
10	1·5	1·1
5	1·7	1·1
2·5	1·9	1·1
1	2·1	1·1
0·1	2·6	1·25
0·01	3·0	1·25
0·005	3·1	1·3

However, the main purpose of the tabulation was to compare the negative binomials with the corresponding G.W. distributions when $\rho = 4{\cdot}5$.

Every percentage point tabulated is smaller in the negative binomials than in the G.W. distribution with $\rho = 4{\cdot}5$, and the same arguments q_a, q_k (except for the 10 per cent point when $q_a = q_k = 0{\cdot}25$; both are then 2). This effect becomes more marked as $\beta/100$, the probability corresponding to the percentage point x_β declines. For example, we may compare the values when $q = q_a = q_k = 0{\cdot}5$ or 0·9.

Percentage point	*G.W.D.* ($\rho = 4{\cdot}5$)		*Negative binomials*	
	$q = 0{\cdot}5$	$q = 0{\cdot}9$	$q = 0{\cdot}5$	$q = 0{\cdot}9$
10	13	814	9	444
5	17	1,031	10	469
2·5	22	1,279	12	492
1	30	1,667	14	519
0·1	59	3,050	18	578
0·01	108	5,350	22	631
0·005	128	6,300	23	645

At $q = 0{\cdot}5$, the ratio of corresponding percentage points (G.W.D./N.B.) varies between 1·4 and 5·6, at $q = 0{\cdot}9$ between 1·8 and 9·8. Intermediate values of q_a, q_k give intermediate results.

At $q = 0{\cdot}25$, the 10 per cent points are both 2 for the G.W.D. and negative binomials; the 5 per cent points are respectively 3 and 2; the 0·005 per cent points are

29 for the G.W.D. and 7 for the negative binomials. The conclusion already reached from Section 6.4, but here reinforced, is that the G.W.D.s with $\rho = 4\cdot5$ have much longer tails than the negative binomials ($\rho \to \infty$) with the same values of k and q_a.

6.742. In Section 6.2, we compared the mean, median and mode for the two sets of cases which give rise to the same value of k.

We now make the same comparison for the percentage points:

(a) (i) $\rho = 1\cdot5, q_k = 0\cdot5$, (ii) $\rho = 4\cdot5, q_k = 0\cdot25$, (iii) negative binomial with $k = 1\cdot5$;

(b) (i) $\rho = 0\cdot5, q_k = 0\cdot9$, (ii) $\rho = 4\cdot5, q_k = 0\cdot5$, (iii) negative binomial with $k = 4\cdot5$.

Since the percentage points increase with q_a, a sufficient summary of the position is provided by comparing the values for $q_a = 0\cdot25$ and $0\cdot9$.

Percentage point	a(i) *G.W.D.* $\rho = 1\cdot5$ $q_k = 0\cdot5$ $q_a = 0\cdot25$	0·9	a(ii) *G.W.D.* $\rho = 4\cdot5$ $q_k = 0\cdot25$ $q_a = 0\cdot25$	0·9	a(iii) *Negative binomial* $k = 1\cdot5$ $q_a = 0\cdot25$	0·9
10	3	73	2	39	2	29
5	6	127	3	54	2	36
2·5	10	211	4	71	3	44
1	19	404	6	98	3	53
0·1	96	1,939	13	195	5	76
0·01	450	9,100	24	357	7	99
0·005	716	14,400	29	424	7	106

Percentage point	b(i) *G.W.D.* $\rho = 0\cdot5$ $q_k = 0\cdot9$ $q_a = 0\cdot25$	0·9	b(ii) *G.W.D.* $\rho = 4\cdot5$ $q_k = 0\cdot5$ $q_a = 0\cdot25$	0·9	b(iii) *Negative binomial* $k = 4\cdot5$ $q_a = 0\cdot25$	0·9
10	30	2,300	5	101	3	67
5	126	9,150	7	132	4	78
2·5	511	36,500	10	168	5	88
1	3,200	238,000	13	224	6	101

The distributions (ii) and (iii) have already been compared in the previous section. The difference between a(i) and a(ii) is shown up much more clearly by the percentage points than by the mean, median and mode (see p. 220). So, *a fortiori* is the difference between b(i) and b(ii); in particular the reduction of ρ to a value below unity altogether outweighs the increase in q_k.

6.75. We now consider the percentage points of the negative binomials in *standard measure*. Table 5 is helpful in relation to the following discussion.

The relative constancy of any percentage point for changing values of q_a, q_k is again a striking feature of the results. As remarked on p. 220, when q_a or $q_k = 0\cdot25$ or $0\cdot5$, the discrete nature of the distributions may produce irregularities in consecutive results. However, it is clear that, in general, each percentage point decreases as q_a increases from $0\cdot5$ to $0\cdot9$ for fixed q_k (or k) and as q_k (or k) increases for fixed q_a.

The decrease may be monotonic; in general this appears to be the case. The percentage point as here defined is the smallest value of the variate for which the tail frequency is less than $\beta/100$. So the "true" percentage point could be anything between x_β and $x_\beta - 1$; or, in standard measure, $(x_\beta - \mu)/\sigma = X_\beta$ and $(x_\beta - \mu - 1)/\sigma = X_\beta - 1/\sigma$. In Table 3 the intervals $(X_\beta - 1/\sigma) - X_\beta$ overlap in all cases for consecutive values of q_a (q_k constant) or q_k (q_a constant) with two exceptions:

(i) Ten per cent point, $q_k = 0{\cdot}25$ ($k = 1{\cdot}5$), $q_a = 0{\cdot}5, 0{\cdot}6$. There is a genuine drop between $q_a = 0{\cdot}5$ and $0{\cdot}6$ with a subsequent rise at $q_a = 0{\cdot}7$. The intervals are (0·8–1·4), (0·3–0·7), (1·0–1·3).

(ii) Five per cent point, $q_k = 0{\cdot}6, q_a = 0{\cdot}25, 0{\cdot}5$. There is a genuine rise from $q_a = 0{\cdot}25$ to $q_a = 0{\cdot}5$. The intervals for $q_a = 0{\cdot}25, 0{\cdot}5, 0{\cdot}6$ are (1·0–1·6), (1·7–2·0), (1·6–1·8).

TABLE 5

Upper percentage points in standard measure for negative binomial $\{1/p_a - (q_a A)/p_a\}^{-k}$ for low and high values of q_a, q_k

			Negative binomials*							
			(i)				(ii)			
		q_a	0·25	0·25	0·25	0·9	0·25	0·5	0·9	0·9
Percentage point	*Normal values*	q_k	0·25	0·5	0·9	0·9	0·25	0·25	0·25	0·9
		k	1·5	4·5	40·5	40·5	1·5	1·5	1·5	40·5
10	1·3		1·8	1·1	1·3	1·3	1·8	1·4	1·3	1·3
5	1·6		1·8	1·8	1·8	1·7	1·8	2·0	1·9	1·7
2·5	2·0		3·1	2·5	2·2	2·1	3·1	2·6	2·5	2·1
1	2·3		3·1	3·2	2·7	2·6	3·1	3·2	3·2	2·6
0·1	3·1		5·5	4·6	3·7	3·5	5·5	5·5	5·2	3·5
0·01	3·7		8·0	6·0	4·6	4·4	8·0	7·2	7·2	4·4
0·005	3·9		8·0	6·7	4·8	4·6	8·0	8·4	8·0	4·6

* $k = 4{\cdot}5 q_k/p_k$, where the argument q_k is the same as in the G.W.D. with $\rho = 4{\cdot}5$.

If the percentage points are taken to be $X_\beta - 1/2\sigma$ the decrease is not always monotonic between q_a or $q_k = 0{\cdot}25$ and $0{\cdot}6$. It is monotonic for the 0·1 and 0·01 per cent points. In the other cases there are exceptions. One such exception shows up in Table 5 where for $q_a = 0{\cdot}25$, $q_k = 0{\cdot}25$, 0·5 and 0·9, the tabulated values are 1·8, 1·1, 1·3 and the intervals (0·6–1·8), (0·4–1·1), (1·1–1·3).

The main interest is, of course, in comparing the negative binomials (e.g. Table 5 or 3) with the G.W.D.s (Table 4 or 3) for $\rho = 4{\cdot}5$ and the same value of q_k (or k). The 10 per cent points are larger in the negative binomials, the 5 per cent points almost the same.

The $2\frac{1}{2}$ per cent points are larger when $q_a = q_k = 0{\cdot}25$ but otherwise smaller. The 1, 0·1, 0·01 and 0·005 per cent points are all markedly smaller in the negative binomials; and this effect increases as the percentage diminishes. The values of the 0·01 per cent points are less than half (but always more than one-fifth) as large in the negative binomials. All percentage points in the negative binomials (ii) are equal to or greater than the corresponding values in (i).

The 10 per cent points of the negative binomials are larger or about the same size as their normal values. All other percentage points are larger than their normal values increasingly so as β diminishes. Thus, in general, the negative binomials have long tails, but not nearly so long as the corresponding G.W.D.s, even at $\rho = 4{\cdot}5$.

7. EXAMPLES OF FITTING THE GENERALIZED WARING DISTRIBUTION

7.1. In a paper written in 1962 (Irwin, 1963) the Simple Waring Distribution was fitted by maximum likelihood to the observed distribution of the number of filarial worms (*Litomosoides carinii*) on 2,600 mites (*Liponyssus bacoti*). Using a computer, Dr Norman Bailey, to whom I am much indebted, has fitted the same data with the Generalized Waring Distribution. He used two methods. (1a) In the first, he fitted by expressing the likelihood in terms of a, k, ρ, obtaining from the computer sufficient values to derive directly the maximum and the corresponding values of a, k. (1b) In the second, he used the fact that all the theoretical frequencies are expressible in terms of $\theta = a+k$ and $\phi = ak$ and calculated the likelihood in terms of θ and ϕ, deriving the maximum and transforming back to a, k afterwards.

Of (1a) he writes:

"Actual mapping of L shows that certain difficulties arise; e.g. for some fixed ρ the surface is bimodal as expected since symmetric in a, k. But for some values there is a single maximum on $a = k$. One has doubts whether the surface is satisfactory for M.L. estimation."

This objection does not apply to the second method. In both cases he obtained standard errors; in (1a) from $\partial^2 L/\partial a^2$, $\partial^2 L/\partial a\,\partial k$, $\partial^2 L/\partial k^2$ directly; in (1b) from $V(a) = \{a^2\,V(\theta) - 2a\,\mathrm{cov}(\theta, \phi) + V(\phi)\}\,|\,(a-k)^2$ with a similar expression for $V(k)$.

The results for (1a) and (1b) agreed closely; but (1b) has much smaller standard errors.

Table 6 shows the observed and expected values of the frequencies, as well as the values of χ^2 for the Simple Waring Distribution and the Generalized Waring Distribution, using the same grouping as in the 1963 paper. The values of $\hat{a}, \hat{k}, \hat{\rho}$ are as follows:

	Simple Waring	*Generalized Waring* (1a)	*Generalized Waring* (1b)
$\hat{\rho}$	1·85	$1{\cdot}68 \pm 0{\cdot}08$	$1{\cdot}67 \pm 0{\cdot}04$
$\hat{a}$	2·35	$1{\cdot}53 \pm 0{\cdot}35$	$1{\cdot}46 \pm 0{\cdot}09$
$\hat{k}$	1	$1{\cdot}31 \pm 0{\cdot}28$	$1{\cdot}37 \pm 0{\cdot}09$
$\chi^2_{25} =$	33·8	$\chi^2_{24} =$ 28·5	28·0
P	0·11	P 0·24	0·26

There is clearly no appreciable difference between (1a) and (1b). The standard errors for the Simple Waring Distribution were not calculated, but the results are clearly different from those given by the G.W.D.; for in the former case $k = 1$ theoretically. The G.W.D. shows a little improvement.

7.2. A second example is Newbold's accident distribution (Newbold, 1925, 1927) to which (Irwin, 1968) the Generalized Waring Distribution was fitted and showed some improvement on Newbold's negative binomial. This is not a particularly long tailed distribution and was fitted by equating three factorial moments. Dr Bailey has fitted this by maximum likelihood, using methods (a) and (b) already described.

The results are as follows:

	Factorial moments	Generalized Waring (2a)	Generalized Waring (2b)
$\hat{\rho}$	7·55	6·85 ± 1·87	6·92 ± 2·23
$\hat{a}$	6·05	4·38 ± 2·02	4·31 ± 1·61
$\hat{k}$	1·06	1·38 ± 0·43	1·33 ± 0·81
χ^2_4	10·6	10·7	10·6
P	0·03	0·03	0·03

TABLE 6

Distribution of the number of filarial worms on 2,600 mites (Bertram's data)
(Observed and expected numbers calculated from fitted Waring Distributions)

No. of worms	Observed	Simple Waring	Expected G.W.D. (1a)	Expected G.W.D. (1b)
0	1,155	1,145·2	1,165·5	1,162·1
1	553	517·6	516·8	516·5
2	265	279·6	273·6	273·8
3	150	169·0	163·4	163·7
4	98	110·2	106·1	106·4
5	70	76·1	73·1	73·4
6	48	54·8	52·7	53·0
7	36	40·9	39·4	39·6
8	28	31·3	30·3	30·5
9	27	24·6	23·9	24·0
10	15	19·6	19·2	19·3
11	13	16·0	15·7	15·8
12	21	13·1	13·0	13·1
13	8	11·0	10·9	11·0
14	11	9·2	9·2	9·3
15	5	7·9	7·9	8·0
16	9	6·8	6·8	6·9
17	9	5·9	5·9	6·0
18	9	5·1	5·2	5·2
19	8	4·5	4·6	4·6
20	5	3·9	4·1	4·1
21	2	3·5	3·6	3·7
22	4	3·1	3·2	3·3
23	5	2·8	2·9	2·9
24	3	2·5	2·6	2·7
25	5	2·3	2·4	2·4
26	1	2·1	2·2	2·2
>26	37	31·4	35·9	36·6
Total	2,600	2,600	2,600·1	2,600·1
χ^2_{25}		33·8	χ^2_{24} 28·5	28·0
P		0·11	0·24	0·26

The observed and expected frequencies are shown in Table 7. Neither (2a) nor (2b) show any significant difference from the fit by factorial moments.

TABLE 7

Accidents to men in a soap factory (5 *months' exposure*)

	Observed		*Negative binomial*	*Generalized Waring*		
				Factorial moments	*Max L(a)*	*Max L(b)*
0	239		251	240	236	237
1	98		93	105	109	108
2	57		46	49	50	50
3	33		25	24	24	24
4	9		14	12	12	12
5	2		8	7	7	7
6	2		4	4	2	4
7	1					
8	—					
9	4					
10	1	7	6	6	7	5
11	—					
12	—					
13	1					
Total	447		447	447	447	447
χ^2			$\chi_5^2 = 13.7$	$\chi_4^2 = 10.6$	10·7	10·6
P			0·018	0·032	0·030	0·032

Whereas in the first example method (b) gave considerably smaller standard errors than (a), here they are on the whole slightly larger. Dr Bailey remarks:

"The inherent accuracy of the estimates is low and we may well doubt whether the standard errors are meaningful."

One may add that they are "large sample" standard errors and compared with the first example the sample is small.

[To be concluded.]

REFERENCES

IRWIN, J. O. (1963). The place of mathematics in medical and biological statistics. *J. R. Statist. Soc.* A, **126**, 1–44.

—— (1968). The generalized Waring distribution, applied to accident theory. *J. R. Statist. Soc.* A, **131**, 205–225.

NEWBOLD, E. M. (1925). A contribution to the study of the human factor in the causation of accidents. *Industrial Health Research Board Report No.* 34.

—— (1927). Applications of the statistics of repeated events, particularly to industrial accidents. *J. R. Statist. Soc.* A, **90**, 487–547.