A STUDY OF THE DISTRIBUTION OF MEANS ESTIMATED FROM SMALL SAMPLES BY THE METHOD OF MAXIMUM LIKELIHOOD FOR PEARSON'S TYPE II CURVE

By John L. Carlson

The object of this paper is to study the distribution of estimates of the parameter of location for Pearson's Type II Curve. estimated by the method of maximum likelihood from small samples.

R. A. Fisher has assumed,[1] and Professor Hotelling has proved[2] that in large categories of cases the distribution of an optimum statistic approaches normality as the sample size increases. This normality has been assumed to hold for optimum statistics in general whether calculated from large samples or small ones, and it has also been assumed that optimum statistics have minimum variance and always give better fits than do statistics calculated by the method of moments. That this is the case whenever the sample is large and the distribution of optimum statistics normal is made plausible by the reasoning of R. A. Fisher.[3] In case the sample is small, however, there may be reason to doubt that the normality of distribution of optimum statistics holds, and that the other conclusions hold. It is with this phase of the subject that we shall be concerned in what follows.

Before entering into the topic under discussion it will be convenient to review some of the more elementary facts regarding the curve with which we are to be concerned. We shall take first the general equation for the curve in the form,

$$y = y_0\left(1 - \frac{(x-m)^2}{a^2}\right)^p \tag{1}$$

[1] On the Mathematical Foundations of Theoretical Statistics, R. A. Fisher, Phil. Trans. Series A, Vol. 222, 1922. Pp. 309-368.

[2] The Consistency and Ultimate Distribution of Optimum Statistics. Harold Hotelling, Trans. Amer. Math. Soc. Vol. 32, No. 4. Pp. 847-859.

[3] Ibid, Pp. 328-368.

and determine the effect of variation of the constants.

When $p=0$ the equation reduces to the straight line $y=y_0$

When $p=+1$ the equation is that of a parabola with $y=y_0$ at the point $x=m$ and the x intercepts at the points $x=\pm a$. It of course meets the x axis at an angle.

When $p=+2$ the equation $y=0$ is of the fourth degree in x with double roots at the points $x=\pm a$.

In general when $p=n$ the equation $y=0$ is of degree $2n$ in x and has two sets of n-fold multiple roots $x=\pm a$.

Since our curve is to be a probability curve it will of necessity have unit area, and this fact makes it possible for us to evaluate y_0 in terms of the parameters a and p. In order to do this we shall perform the integration below.

$$\text{Area} = 1 = \int_{x-a}^{x+a} y\,dx = 2y_0\int_0^a \left(1-\frac{x^2}{a^2}\right)^p dx$$

whence

$$1 = 2ay_0\int_0^{\frac{\pi}{2}} \cos^{2p+1}\theta\,d\theta$$

but now since

$$(2) \qquad \int_0^{\frac{\pi}{2}} \cos^n\theta\,d\theta = \frac{\sqrt{\pi}}{2}\,\frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)}$$

we have $$1 = 2ay_0\frac{\sqrt{\pi}}{2}\,\frac{\Gamma(p+1)}{\Gamma(p+\frac{3}{2})}$$

and so

$$y_0 = \frac{\Gamma(p+\frac{3}{2})}{a\sqrt{\pi}\,\Gamma(p+1)}$$

Therefore

$$(3) \qquad y = \frac{\Gamma(p+\frac{3}{2})}{a\sqrt{\pi}\,\Gamma(p+1)}\left[1-\frac{(x-m)^2}{a^2}\right]^p$$

Formula (2) we shall find to be of value a number of times and formula (3) is the form in which equation (1) will be used throughout the remainder of the paper.

It will be worth while now to consider the likelihood function L together with its first and second partial derivatives with respect to m. (We shall hereafter refer to m as the parameter of location, a as the parameter of scaling, and p as the parameter of shape.) We are to use m to denote the estimate of m obtained by the method of maximum likelihood, in accordance with the convention introduced by Fisher,[1] and it will be with this parameter that we shall concern ourselves in this investigation.

We have from (3) on the preceding page

$$(4) \qquad L = n \log \frac{\Gamma(p+\frac{3}{2})}{a\sqrt{\pi}\,\Gamma(p+1)} + p\sum_{i=1}^{n} \log\left[1-\frac{(x_i-m)^2}{a^2}\right]$$

and so

$$(5) \qquad \frac{\partial L}{\partial m} = 2p\sum_{i=1}^{n} \frac{x_i-m}{a^2-(x_i-m)^2}$$

and

$$(6) \qquad \frac{\partial^2 L}{\partial m^2} = -2p\sum_{i=1}^{n} \frac{a^2+(x_i-m)^2}{[a^2-(x_i-m)^2]^2}$$

[1]Ibid Pp. 309-368.

At this point we shall stop to consider the effect of variation of the parameters a and p upon our estimate of $\hat{m}$. Let us first consider p. Since the method of maximum likelihood is here merely the method of the differential calculus it follows from a consideration of equation (5) that our estimate of $\hat{m}$ will be independent of p. for any particular sample. Such is not the case when we consider a. however, for any change in a allows a change in the variance of $\hat{m}$ for the particular sample.

We shall find it advantageous to cover as much of the theoretical work as possible before embarking upon our experimental check and its great amount of numeral calculations, for it is only by means of a check between theory and experiment that we are able in the present state of knowledge, to judge of the applicability of the method of maximum likelihood to small samples. Of course it will be necessary to consider the distribution of our estimates of $\hat{m}$, in order to make this statistic of practical use. and it is desirable to know the theoretical variances of $\bar{x}$ the arithmetic mean, $\hat{m}$, and the experimentally obtained variance of the distribution of our estimates of $\hat{m}$. The first of these we can obtain from theory, the second from an approximation valid only in the limit, and the last by means of calculation based on actual sampling.

We shall be concerned first with the theoretical variance of $\bar{x}$. It is well known that the variance of $\bar{x}$ is equal to the variance of the distribution divided by n. We must first, therefore, find the variance of the distribution.

From (3) page 88 it follows that

$$\sigma^2 = \frac{2\Gamma(p+\frac{3}{2})}{a\sqrt{\pi}\,\Gamma(p+1)} \int_0^a x^2\left(1-\frac{x^2}{a^2}\right)^p dx$$

$$\sigma^2 = \frac{2\Gamma(p+\frac{3}{2})}{a\sqrt{\pi}\,\Gamma(p+1)} \int_0^{\frac{\pi}{2}} \left[\cos^{2p+1}\theta - \cos^{2p+3}\theta\right] d\theta$$

remembering now (2) page 87

$$\sigma^2 = \frac{a^2}{2p+3}$$

whence it follows that

(7) $$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} = \frac{a^2}{n(2p+3)}.$$

We shall now calculate the limiting form of the variance of $\hat{m}$. Fisher has proved[1] that if the distribution of optimum statistics is normal the variance of an optimum statistic is equal to the negative reciprocal of the mathematical expectation of the second partial derivative of the logarithm of the likelihood with respect to the parameter in question.

We may write, therefore

$$-\frac{1}{\sigma_{\hat{m}}^2} = \frac{-4np\Gamma(p+\frac{3}{2})}{a\sqrt{\pi}\,\Gamma(p+1)}\int_0^a \frac{(a^2+x^2)}{(a^2-x^2)}\left[1-\frac{x^2}{a^2}\right]^p dx$$

whence

$$-\frac{1}{\sigma_{\hat{m}}^2} = \frac{-4np\Gamma(p+\frac{3}{2})}{a\sqrt{\pi}\,\Gamma(p+1)}\int_0^{\frac{\pi}{2}} (2\cos^{2p-3}\theta - \cos^{2p-1}\theta)\,d\theta$$

and so again referring to (2) page 87 we have

$$-\frac{1}{\sigma_{\hat{m}}^2} = -\frac{2np(p+\frac{1}{2})}{a^2(p-1)}$$

hence we have

(8) $$\sigma_{\hat{m}}^2 = \frac{a^2(p-1)}{np(2p+1)}$$

The efficiency of the mean is then

(9) $$E = \frac{(p-1)(2p+3)}{p(2p+1)} = \frac{\sigma_{\hat{m}}^2}{\sigma_{\bar{x}}^2}$$

The next problem with which we must concern ourselves is that of experimental verification of the assumptions under discus-

[1] Ibid Pp. 327-328.

sion. The problem is briefly that of choosing a number of small samples from a population which obeys our law of frequency, estimating m for these samples, and then calculating the variance for the distribution. Also we shall draw an histogram of the distribution and observe the general type of the distribution, so nearly as that is possible from our samples.

The problem of choosing our samples is not the least of our difficulties, for we can not take all types of samples. They must be of a very special nature: they must be from a population of the Type II. In order to accomplish this it will be necessary to have a table of areas corresponding to given values of x for the Type II Curve. There is no such table available to the knowledge of the writer, and it is therefore necessary to construct the table before we can procede with the choosing of the samples. After the table has been built, we can with the aid of Tippett's Tables,[1] choose our samples with ease. The manner of choosing is as follows. Take the numbers from Tippett's Tables as areas under the Type II Curve and look up in the table of areas the values of x corresponding to the smallest area containing the area found from the Random Numbers. This will give the value of x to be taken. Since we will take four digits let the fifth digit determine the sign. If it is odd take the sign $-$, if it is even take the sign $+$.

There are two ways in which a table of this nature can be prepared, and the method employed must in any case be determined by the degree of accuracy attainable and the amount of labor involved. One of these methods is that of the calculus of finite differences, determining the zero order differences by means of algebra, and from these by the process of addition building up the table. This method is best used when a dependable listing adding machine is at hand. The other method is that of direct integration, and it is found that with the aid of two calculating machines this is by far the quicker. It was this method that was applied in the building of the table on page 92 and 93.

[1]Tracts For Computers, No. XV.

Table of Areas Under Pearson's Type II Curve, Correct to 9 Places of Decimals. The Areas are included between ordinates located $\pm x$ units from the parameter of location.

Constants. $y_0 = {}^{15}/_{16}\,a$, $m = 0$, $p = 2$, and $a = 1$

x	Area	x	Area
0.00	0.000 000 000	30	529 661 250
01	018 748 750	31	545 084 843
02	037 490 001	32	560 298 291
03	056 212 259	33	575 295 077
04	074 920 038	34	590 073 828
05	093 593 867	35	604 625 820
06	112 230 292	36	618 947 482
07	130 821 880	37	633 034 148
08	149 361 229	38	646 881 319
09	167 840 964	39	660 484 657
10	186 253 750	40	673 840 000
11	204 592 289	41	686 943 358
12	222 849 331	42	699 790 921
13	241 017 673	43	712 379 067
14	259 090 168	44	724 704 358
15	277 059 727	45	736 763 555
16	294 919 322	46	748 553 612
17	312 661 995	47	760 071 688
18	330 280 859	48	773 151 488
19	347 769 104	49	782 281 572
20	365 120 040	50	792 968 750
21	382 326 904	51	803 374 696
22	399 383 261	52	813 497 651
23	416 282 613	53	823 336 081
24	433 018 598	54	832 888 688
25	449 584 961	55	842 154 414
26	465 975 552	56	851 132 442
27	482 184 334	57	860 009 702
28	498 205 389	58	868 223 379
29	514 032 918	59	876 335 911
30	529 661 250	60	884 160 000

61	892 546 290	81	985 203 165
62	898 944 981	82	987 317 441
63	905 907 620	83	989 230 274
64	912 585 318	84	990 949 478
65	919 120 273	85	992 483 242
66	925 092 472	86	993 840 132
67	930 925 941	87	995 029 095
68	936 482 509	88	996 059 469
69	941 764 926	89	996 940 979
70	946 776 250	90	997 683 750
71	951 519 851	91	998 298 304
72	955 999 411	92	998 795 571
73	960 294 310	93	999 186 888
74	964 182 748	94	999 484 008
75	967 895 508	95	999 699 102
76	971 362 202	96	999 844 762
77	974 588 156	97	999 934 010
78	977 579 039	98	999 980 299
79	980 340 865	99	999 997 519
80	982 880 000	100	1.000 000 000

The table on the preceding page was built by direct integration of (3) page 88, with $p=2$ and $a=1$ These values for the parameters were chosen so as to save as much labor in calculation as possible, and at the same time maintain the desired shape of the curve. There has of course been no less of generality in setting $a=1$ but we have limited ourselves quite definitely in using the value $p=2$. The accuracy of the table to 9 places of decimals has been assured by calculating all values to 13 places and then determining the maximum error over the whole range which was 625 in the 13 place due to the use of the decimal equivalent of $^2/_3$ in the third degree term.

Our table of areas being complete and the problem of sampling thus solved, we must next consider the task of estimating $\hat{m}$. Since we have already taken $a=1$ and $p=2$ in building our tables we shall continue to use these values throughout the work.

The next question to be settled before beginning our work is the size of the samples with which we are to deal. The case $n = 1$ is of course of no interest for the best estimate of a single observation is the observation itself. The case $n = 2$ is likewise of no interest for in this case the arithmetic mean coincides with the solution by the method of maximum likelihood. Therefore it is the case $n = 3$ with which we shall be concerned. Our results are not then trivial, and at the same time we have the case which is the easiest to deal with, since the number of numerical calculations for each sample is reduced to the minimum.

Before going ahead with any attempt at solving an equation such as (5) page 88, it is well to have in mind a picture of what we are actually trying to accomplish. With such a picture in mind we are better able to realize the difficulties of the situation and so are better able to cope with them. To this end we have included a graph of the problem involved which will make clear at a glance just what must be done to find the true value of m.

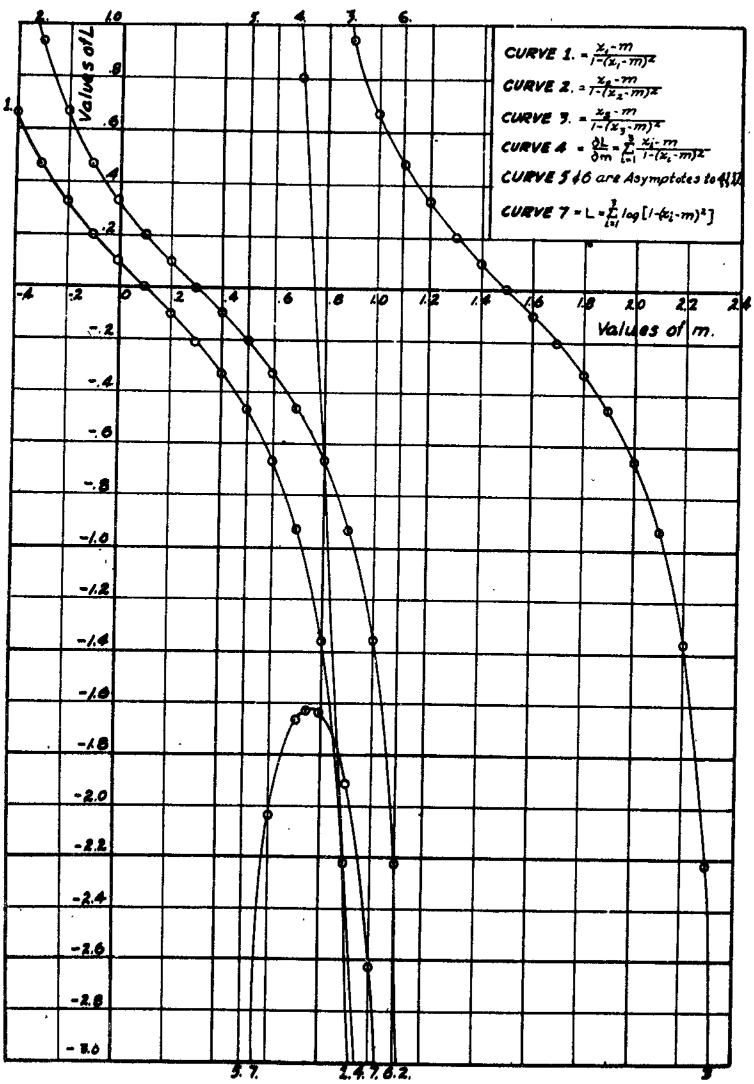
We have drawn, page 96, from plotted points the curves representing L, $\frac{\partial L}{\partial m}$ and the three terms of $\frac{\partial L}{\partial m}$. Also we have drawn in the asymptotes to the curve which are of significance. If we are able to find the point at which $\frac{\partial L}{\partial m} = 0$ we have the solution to our problem. This involves solving a fifth degree equation. We shall use Newton's method of successive approximation. Fisher states[1] that in some cases at least, we may start with an inefficient statistic and by a single approximation obtain an efficient one. Whether or not this is the case for small samples we shall see when our calculations have been analyzed.

It will be seen upon examining the graph that as is allowed to vary each of the therms of $\frac{\partial L}{\partial m}$ varies from $-\infty$ to $+\infty$,

[1]Theory of Statistical Estimation, R. A. Fisher, Proc. Cam. Phil. Soc. Vol. XXII part 5. Pp. 708-709.

and so their sum also varies between these limits. It will be seen that the asymptote corresponding to the largest allowable value of m can be found by adding the value of a to the smallest observation, and that the asymptote corresponding to the smallest allowable value of m may be found by subtracting the value of a from the greatest observation. It is thus evident that as dispersion in the sample increases the variance of m must surely decrease. That this fact is of fundamental importance in choosing our first estimate of m will be seen from consideration of the following case, since it is well known that in the case of a curve with a real finite pole Newton's method may lead us to erroneous results. Let us consider the sample consisting of the three observations $x_1 = -.99$, $x_2 = +.99$, $x_3 = +.99$, for which the arithmetic mean is $\bar{x} = +.33$. If now we take as our first approximation $m_1 = \bar{x}$ we will immediately lead ourselves to an erroneous value of m_2, our estimate of $\hat{m}$. That this is the case will be easily seen by means of the check given on page 94 for using $a = 1$ we locate the asymptotes at $x = \pm .01$. The true value of m located between these asymptotes, is not therefore even to be approached should we take $m_1 = \bar{x}$. This is an extreme case and fortunately not to be anticipated very often, or the arithmetic mean would be deprived of any value whatsoever. We should always be sure that the difference between any observation and the value of m that we are using is not greater than the value of a when dealing with the Type II Curve. This holds no matter what our manner of attack may be.

We shall now be concerned with the calculations for our 100 samples of 3. All of the data are tabulated in such a way as to be self-explanatory, and so we shall not bother to give sample calculations. The next ten pages cover these calculations. The discussion is continued on page 104.


CURVE 1. = $\frac{x_1 - m}{1-(x_1-m)^2}$
CURVE 2. = $\frac{x_2 - m}{1-(x_2-m)^2}$
CURVE 3. = $\frac{x_3 - m}{1-(x_3-m)^2}$
CURVE 4 = $\frac{\partial L}{\partial m} = \sum_{i=1}^{3} \frac{x_i - m}{1-(x_i-m)^2}$
CURVE 5 & 6 are Asymptotes to 4
CURVE 7 = $L = \sum_{i=1}^{3} \log[1-(x_i-m)^2]$
Values of L
Values of m.

Sample No	i	x_i	$m_1=\bar{x}$ $\frac{1}{3}\sum_{i=1}^{3} x_i$	$\frac{x_i-m}{1-(x_i-m)^2}$	$\frac{\partial L}{\partial m}$ $\sum_{i=1}^{3}\frac{x_i-m}{1-(x_i-m)^2}$	$-\frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	$\frac{\partial^2 L}{\partial m^2}$ $-\sum_{i=1}^{3}\frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	m_2 $\frac{\frac{\partial^2 L}{\partial m^2}m_1-\frac{\partial L}{\partial m}}{\partial^2 L/\partial m^2}$
	1	+.17		+ .254667		- 1.120600		
1	2	-.23	-.0700	- .164200	+.009840	- 1.051500	- 3.184900	-.066879
	3	-.15		- .080520		- 1.012800		
	1	-.03		- .083882		- 1.021059		
2	2	-.15	+.0533	- .212064	+.016730	- 1.133055	- 3.438830	+.049737
	3	+.34		+ .312676		- 1.284716		
	1	+.67		+ .254668		- 1.190833		
3	2	-.11	+.4300	- .762281	-.177943	-2.573777	- 5.080876	+.465022
	3	+.73		+ .329670		- 1.316266		
	1	-.54		- .291230		- 1.248262		
4	2	-.33	-.2700	-.060216	+.018882	-1.010865	- 3.655622	-.264835
	3	+.06		+.370328		-1.396495		
	1	-.41		-.481688		-1.718538		
5	2	+.11	+.0067	+.104445	-.029803	-1.032609	- 4.101439	+.013936
	3	+.32		+.347440		-1.350292		
	1	-.26		-.408701		-1.479856		
6	2	+.12	+.0967	+.023312	-.024353	-1.001630	- 3.858903	+.103011
	3	+.42		+.361036		-1.377417		
	1	+.54		+1.770451		-4.424677		
7	2	-.50	-.2167	- .308021	+.853498	-4.277016	- 9.978709	-.131268
	3	-.69		- .609932		-1.277016		
	1	+.19		+.815850		-2.218101		
8	2	-.64	-.3700	- .291230	+.207992	- 1.248262	- 4.758692	-.326292
	3	-.66		- .316628		- 1.292329		
	1	-.12		+ .361036		- 1.377433		
9	2	-.48	-.4433	- .036749	+.011911	- 1.004049	- 3.666198	-.440051
	3	-.73		- .312376		- 1.284716		
	1	-.49		+ .167774		- 1.083693		
10	2	-.77	-.6533	- .118311	+.002661	- 1.041802	- 3.132061	-.652450
	3	-.70		- .046802		- 1.006566		
	1	-.07		+.066896		-1.013446		
11	2	+.04	-.1367	+.182285	-.009546	-1.098764	- 3.308885	-.139555
	3	-.38		- .258727		-1.196675		
	1	-.40		+ .128735		- 1.049476		
12	2	-.72	-.5267	- .200836	-.005134	-1.119217	- 3.182127	-.528283
	3	-.46		+ .066967		- 1.013434		
	1	-.30		- .036749		- 1.004049		
13	2	-.65	-.2633	- .454723	-.387058	-1.589322	- 4.343556	-.352411
	3	+.16		+ .104414		-1.750185		
	1	-.14		+ .128767		-1.049476		
14	2	-.50	-.2667	- .246729	-.010034	-1.179312	- 3.263601	-.269775
	3	-.16		+ .107928		-1.034813		

Sample No	i	x_i	$m_1 = \bar{x}$ $\frac{1}{3}\sum_{i=1}^{3} x_i$	$\frac{x_i - m}{1-(x_i-m)^2}$	$\frac{\partial L}{\partial m}$ $\sum_{i=1}^{3} \frac{x_i - m}{1-(x_i-m)^2}$	$\frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	$\frac{\partial^2 L}{\partial m^2}$ $\sum_{i=1}^{3} - \frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	m_2 $\frac{\frac{\partial^2 L}{\partial m^2} m_1 - \frac{\partial L}{\partial m}}{\partial^2 L/\partial m^2}$
	1	-.09		+ .149926		- 1.066950		
15	2	-.20	-.2367	+ .036749	+ .002713	- 1.004049	- 3.177717	-.235846
	3	-.43		- .183962		- 1.106718		
	1	-.29		- .360623		- 1.376579		
16	2	+.12	+.0333	+ .087663	- .021856	- 1.022996	- 3.585193	+.039096
	3	+.27		+ .251104		- 1.185618		
	1	-.10		+ .142798		- 1.060774		
17	2	-.39	-.2400	- .153452	- .000653	- 1.070113	- 3.131187	-.240209
	3	-.23		+ .010001		- 1.000300		
	1	+.26		+ .428687		- 1.526159		
18	2	-.42	-.1100	- .342958	+ .025513	- 1.341557	- 3.878581	-.103422
	3	-.17		- .060216		- 1.010865		
	1	+.23		+ .423670		- 1.514353		
19	2	-.21	-.1367	- .073695	+ .029076	- 1.016264	- 3.830699	- .129111
	3	-.43		- .320905		- 1.300082		
	1	+.37		+ .053552		- 1.008563		
20	2	-.09	+.3167	- .483174	- .025787	- 1.673112	- 4.150227	+ .322813
	3	+.67		+ .403835		- 1.468552		
	1	+.21		+ .178665		- 1.094805		
21	2	-.76	+.0367	- 2.181131	- .983165	-12.252377	-17.060464	- .094328
	3	+.66		+ 1.019301		- 3.713282		
	1	-.20		- .139303		- 1.057853		
22	2	+.23	-.0633	+ .320905	+ .020958	- 1.300082	- 3.434721	- .057198
	3	-.22		- .160644		- 1.076786		
	1	-.44		- .142798		- 1.060744		
23	2	-.62	-.3000	- .356506	+ .084156	- 1.368375	- 4.378292	-.280779
	3	+.16		+ .583460		- 1.949243		
	1	-.28		- .481639		- 1.658197		
24	2	+.20	+.1233	+ .077153	- .038750	- 1.017823	- 4.063031	- .132837
	3	+.45		+ .365736		- 1.387011		
	1	-.13		+ .107928		- 1.034813		
25	2	-.56	-.2367	- .361036	+ .155433	- 1.377417	- 3.564903	-.193096
	3	-.02		+ .408551		- 1.152673		
	1	+.10		+ .111347		- 1.024632		
26	2	+.14	-.0100	+ .153452	- .014051	- 1.070113	- 3.322761	-.014229
	3	-.27		- .278850		- 1.228016		
	1	-.79		- .603260		- 2.011377		
27	2	-.09	-.3200	+ .242846	- .105746	- 1.142660	- 4.344870	-.344338
	3	-.08		+ .254668		- 1.190833		
	1	-.08		- .273953		- 1.028078		
28	2	+.43	+.1760	+ .271517	+ .001564	- 1.216408	- 3.244534	-.175518
	3	+.18		+ .004000		- 1.000048		

Sample No.	i	x_i	$m_1 = \bar{x}$ $\frac{1}{3}\sum_{i=1}^{3} x_i$	$\frac{x_i - m}{1-(x_i-m)^2}$	$\frac{\partial L}{\partial m}$ $\sum_{i=1}^{3}\frac{x_i-m}{1-(x_i-m)^2}$	$-\frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	$\frac{\partial^2 L}{\partial m^2}$ $\sum_{i=1}^{3} -\frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	m_2 $\frac{\frac{\partial^2 L}{\partial m^2} m_1 - \frac{\partial L}{\partial m}}{\partial^2 L/\partial m^2}$
	1	+.50		+ .864077		- 2.991572		
29	2	-.59	-.0767	- .696923	+.103600	- 2.329134	- 6.332807	-.060341
	3	-.14		- .063554		- 1.012101		
	1	+.02		- .203830		- 1.123044		
30	2	+.17	+.2160	- .046097	+.009519	- 1.006379	- 3.327343	+.213139
	3	+.46		+ .259446		- 1.197929		
	1	-.35		- .057185		- 1.009800		
31	2	-.18	-.2930	+ .114461	+.000091	- 1.039137	- 3.058737	-.292970
	3	-.35		- .057185		- 1.009800		
	1	+.29		+ .329670		- 1.316266		
32	2	-.16	-.0100	- .153452	+.022766	- 1.070113	- 3.456492	-.003414
	3	-.16		- .153452		- 1.070113		
	1	-.45		+ .017004		- 1.000867		
33	2	-.60	-.4667	- .135395	-.000222	- 1.054671	- 3.097560	-.463667
	3	-.35		+ .118623		- 1.042022		
	1	-.23		- .160644		- 1.076786		
34	2	+.21	-.0733	+ .308021	+.018610	- 1.057853	- 3.184115	-.067455
	3	-.20		- .128767		- 1.049476		
	1	-.43		- .444132		- 1.563278		
35	2	-.05	-.0500	+ .000000	+.000000	- 0.000000	- 2.832959	-.050000
	3	+.33		+ .444132		- 1.563278		
	1	-.62		- .630593		- 2.100061		
36	2	-.59	-.1367	- .570533	+6.439597	- 1.909640	-128.928063	-.086753
	3	+.80		+7.640723		-124.918362		
	1	-.44		- .509956		- 1.734292		
37	2	+.50	-.0200	+ .712719	+.101753	- 2.386551	- 5.151350	-.000247
	3	-.12		- .101010		- 1.030507		
	1	-.14		- .312376		- 1.284716		
38	2	+.49	+.1467	+ .387806	+.018548	- 1.436499	- 3.730911	+.130837
	3	+.09		- .056882		- 1.009696		
	1	-.32		- .689282		- 2.301754		
39	2	+.74	+.1900	+ .788530	+.059184	- 2.677252	- 5.983818	+.179792
	3	+.15		- .040064		- 1.004812		
	1	-.17		+ .013002		- 1.000507		
40	2	+.08	-.1833	+ .285611	-.001407	- 1.233970	- 3.497606	-.183402
	3	-.46		- .300020		- 1.263129		
	1	-.30		- .325339		- 1.308220		
41	2	+.07	-.0033	+ .073695	-.016626	- 1.016264	- 3.487431	-.008067
	3	+.22		+ .235018		- 1.162947		
	1	+.01		+ .167774		- 1.909102		
42	2	-.10	-.1533	+ .053451	-.006152	- 1.008563	- 4.070338	-.154811
	3	-.37		- .227577		- 1.153673		

Sample No.	i	x_i	$m_1 = \bar{x}$ $\frac{1}{3}\sum_{i=1}^{3} x_i$	$\frac{x_i - m}{1-(x_i-m)^2}$	$\frac{\partial L}{\partial m}$ $\sum_{i=1}^{3}\frac{x_i-m}{1-(x_i-m)^2}$	$-\frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	$\frac{\partial^2 L}{\partial m^2}$ $\sum_{i=1}^{3} -\frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	m_2 $\frac{\frac{\partial^2 L}{\partial m^2} m_1 - \frac{\partial L}{\partial m}}{\partial^2 L/\partial m^2}$
	1	+.27		+.283036		-1.234771		
43	2	+.03	+.0067	+.023412	+.010962	-1.001644	-3.549514	+.003512
	3	-.28		-.295486		-1.313099		
	1	-.28		-.070344		-1.014820		
44	2	-.13	-.2100	+.080515	+.000170	-1.019406	-3.034526	-.209944
	3	-.22		-.010001		-1.000300		
	1	-.09		-.146304		-1.063775		
45	2	+.19	+.0533	+.139303	+.000301	-1.057853	-3.121762	+.053396
	3	+.06		+.006700		-1.000134		
	1	-.08		+.104414		-1.032590		
46	2	-.28	-.1833	-.097612	+.000102	-1.028495	-3.061219	-.183267
	3	-.19		-.006700		-1.000134		
	1	-.10		+.212064		-1.133055		
47	2	-.70	-.3033	-.470788	-.057922	-1.630045	-3.882558	-.318219
	3	-.11		+.200802		-1.119458		
	1	+.67		+.745508		-2.509221		
48	2	+.07	+.1367	-.066896	+.082154	-1.013405	-5.512458	+.121697
	3	-.33		-.596458		-1.989832		
	1	-.28		-.077153		-1.017823		
49	2	-.15	-.2033	+.053466	-.000381	-1.008563	-3.028016	-.203426
	3	-.18		+.023312		-1.001630		
	1	-.33		-.320905		-1.300082		
50	2	+.45	-.0367	+.637773	+.116066	-2.123915	-4.543455	-.011154
	3	-.23		-.200802		-1.119458		
	1	+.44		+.883184		-3.074927		
51	2	-.79	-.1433	-1.112848	-.166413	-4.196875	-9.291723	-.160910
	3	-.08		+.063251		-2.019971		
	1	-.15		-.010001		-1.000300		
52	2	+.36	-.1400	+.666667	+.011845	-2.222222	-5.270074	-.137752
	3	-.63		-.644821		-2.147552		
	1	+.09		-.208333		-1.128472		
53	2	+.31	+.2900	+.020008	-.014143	-1.001200	-3.232369	+.290044
	3	+.47		+.186027		-1.102697		
	1	+.21		+.114773		-1.039349		
54	2	-.29	+.0967	-.454693	-.044559	-1.589382	-3.883929	+.108173
	3	+.37		+.295361		-1.255198		
	1	+.39		+.186027		-1.102697		
55	2	-.21	+.2100	-.509956	-.069261	-2.600554	-4.894084	+.224152
	3	+.45		+.254668		-1.190833		
	1	-.30		+.097612		-1.028495		
56	2	-.22	-.3967	+.182394	-.015355	-1.098764	-3.382457	-.401240
	3	-.67		-.295361		-1.255198		

Sample No.	i	x_i	$m_1=\bar{x}$ $\frac{1}{3}\sum_{i=1}^{3} x_i$	$\frac{x_i-m}{1-(x_i-m)^2}$	$\frac{\partial L}{\partial m}$ $\sum_{i=1}^{3}\frac{x_i-m}{1-(x_i-m)^2}$	$\frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	$\frac{\partial^2 L}{\partial m^2}$ $\sum_{i=1}^{3}\frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	m_2 $\frac{\frac{\partial^2 L}{\partial m^2}m_1-\frac{\partial L}{\partial m}}{\partial^2 L/\partial m^2}$
	1	-.41		- .325339		- 1.324459		
57	2	-.71	-.1133	- .926626	-.810074	- 3.371716	-48.783525	-.129916
	3	+.78		+4.418971		-44.057350		
	1	+.04		+ .258608		- 1.196675		
58	2	-.69	-.2033	- .637773	-.120557	- 2.123915	- 4.517265	-.229988
	3	+.04		+ .258608		- 1.196675		
	1	+.43		+ .564263		- 1.890704		
59	2	-.48	-.0200	- .583460	-.009196	- 1.949243	- 4.840247	-.021900
	3	-.01		+ .010001		- 1.000300		
	1	+.69		+ .779753		- 2.643417		
60	2	-.17	+.1433	- .347399	+.185625	- 1.350213	- 5.172942	+.107416
	3	-.09		- .246729		- 1.179312		
	1	+.12		+ .164203		- 1.080198		
61	2	+.09	-.0400	+ .132234	-.020191	- 1.052162	- 3.316004	-.046089
	3	-.33		- .316628		- 1.183644		
	1	+.63		+ .308021		- 1.277016		
62	2	+.36	+.3467	+ .013302	-.004016	- 1.000530	- 3.585766	+.347820
	3	+.05		- .325339		- 1.308220		
	1	-.00		- .097509		- 1.028435		
63	2	-.40	+.0967	- .659155	+.163267	- 1.654673	- 4.770116	+.062373
	3	+.69		+ .915915		- 2.087008		
	1	+.13		+ .073695		- 1.016264		
64	2	-.69	+.0567	-1.057383	+.247951	- 7.938848	-13.818279	+.038756
	3	+.73		+1.231645		- 4.863167		
	1	-.39		- .070344		- 1.014820		
65	2	-.13	-.3200	+ .197115	+.005018	- 1.115161	- 3.174239	-.318419
	3	-.44		- .121753		- 1.044258		
	1	-.12		- .394067		- 1.447201		
66	2	+.51	+.2267	+ .308021	-.022492	- 1.255198	- 3.714500	+.232755
	3	+.29		+ .063554		- 1.012101		
	1	-.16		- .193442		- 1.110956		
67	2	+.22	+.0267	+ .200802	+.000660	- 1.119458	- 3.230548	+.026496
	3	+.02		- .006700		- 1.000134		
	1	-.51		- .129081		- 1.049717		
68	2	-.44	-.3833	- .057185	+.003074	- 1.009800	- 3.165866	-.382029
	3	-.20		+ .189340		1.106349		
	1	+.35		+ .036749		- 1.004049		
69	2	+.43	+.3133	+ .118311	-.001929	- 1.041802	- 3.119208	+.321050
	3	+.16		- .156989		- 1.073357		
	1	+.59		+ .379751		- 1.416284		
70	2	+.03	+.2533	- .235018	+.029960	- 1.162947	- 3.618580	+.245021
	3	+.14		- .114773		- 1.039349		

Sample No	i	x_i	$m_1 = \bar{x}$ $\frac{1}{3}\sum_{i=1}^{3} x_i$	$\frac{x_i - m}{1-(x_i-m)^2}$	$\frac{\partial L}{\partial m}$ $\sum_{i=1}^{3} \frac{x_i - m}{1-(x_i-m)^2}$	$\frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	$\frac{\partial^2 L}{\partial m^2}$ $\sum_{i=1}^{3} \frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	m_2 $\frac{\frac{\partial^2 L}{\partial m^2} m_1 - \frac{\partial L}{\partial m}}{\partial^2 L/\partial m^2}$
	1	-.19		+ .040064		- 1.004812		
71	2	-.38	-.2300	- .153452	- .002041	- 1.070113	- 3.111969	-.230656
	3	-.12		+ .111347		- 1.037044		
	1	-.45		- .307638		- 1.276344		
72	2	-.16	-.1667	+ .007000	- .000618	- 1.000147	- 3.539620	-.167175
	3	+.11		+ .300020		- 1.263129		
	1	+.71		+1.155256		- 4.424677		
73	2	-.48	+.0533	- .745580	+.284473	- 2.508265	- 7.979731	+.017651
	3	-.07		- .125203		- 1.046789		
	1	-.33		- .223467		- 1.147540		
74	2	+.15	-.1167	+ .287122	+ .010204	- 1.251454	- 3.407557	-.113705
	3	-.17		- .053451		- 1.008563		
	1	-.19		- .073695		- 1.016264		
75	2	-.39	-.1167	- .295361	+ .025011	- 1.255198	- 2.718663	-.109974
	3	+.23		+ .394067		- 1.447201		
	1	-.01		+ .365736		- 1.387011		
76	2	-.73	- .3367	- .465270	- .057547	- 1.615943	- 4.016400	-.351028
	3	-.27		+ .066998		- 1.013446		
	1	-.63		- .820607		- 2.707683		
77	2	-.09	- .0683	- .021710	+ .089560	- 1.001413	- 7.103430	-.055692
	3	+.53		+ .931877		- 3.294334		
	1	-.04		+ .020008		- 1.001200		
78	2	+.47	- .0600	+ .737032	- .058810	- 2.477060	- 6.266361	-.069385
	3	-.61		- .815850		- 2.788101		
	1	+.39		+ .208333		- 1.128472		
79	2	+.38	+.1900	+ .197115	+ .314714	- 1.115161	- 3.268264	+.093706
	3	+.10		- .090734		- 1.024631		
	1	-.77		- .689282		- 2.301754		
80	2	-.16	-.2600	+ .101010	- .095425	- 1.030507	- 5.020126	-.279008
	3	+.15		+ .492847		- 1.687865		
	1	-.03		+ .204504		- 1.123849		
81	2	-.01	-.2267	+ .227262	- .066846	- 1.152521	- 3.979725	-.243397
	3	-.64		- .498612		- 1.703355		
	1	+.21		- .278850		- 1.228016		
82	2	+.71	+.4700	+ .254668	- .004174	- 1.190833	- 3.420049	+.364611
	3	+.49		+ .020008		- 1.001200		
	1	-.26		- .053451		- 1.008563		
83	2	-.28	- .2067	- .073695	+ .001621	- 1.016264	- 3.074303	-.206173
	3	-.08		+ .128767		- 1.049476		
	1	-.16		- .066998		- 1.013446		
84	2	+.23	-.0933	+ .361036	+ .019230	- 1.057853	- 3.292881	-.087460
	3	-.35		- .274808		- 1.221582		
	1	-.31		- .476190		- 1.643990		
85	2	+.06	+.0900	- .030027	+ .021325	- 1.002704	- 4.430139	+.085186
	3	+.52		+ .527542		- 1.783445		

Sample No.	i	x_i	$m_1=\bar{x}$ $\frac{1}{3}\sum_{i=1}^{3} x_i$	$\frac{x_i-m}{1-(x_i-m)^2}$	$\frac{\partial L}{\partial m}$ $\sum_{i=1}^{3}\frac{x_i-m}{1-(x_i-m)^2}$	$\frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	$\frac{\partial^2 L}{\partial m^2}$ $\sum_{i=1}^{3}-\frac{1+(x_i-m)^2}{[1-(x_i-m)^2]^2}$	m_2 $\frac{\frac{\partial^2 L}{\partial m^2}m_1-\frac{\partial L}{\partial m}}{\partial^2 L/\partial m^2}$
	1	+.36		+ .250748		− 1.185102		
86	2	+.36	+.1233	+ .250748	− .108436	− 1.185102	− 3.533041	+.153992
	3	−.35		− .609932		− 1.162837		
	1	+.56		+1.473657		− 6.399500		
87	2	−.60	−.1567	− .551721	+ .626575	− 1.853371	− 9.508069	−.090801
	3	−.43		− .295361		− 1.255198		
	1	−.45		− .287122		− 1.241454		
88	2	−.29	−.1833	− .107928	+ .038693	− 1.034813	− 3.814450	−.173156
	3	+.19		+ .433743		− 1.538183		
	1	−.09		− .010001		− 1.000300		
89	2	−.39	−.0800	− .341851	+ .005768	− 1.341559	− 3.710134	−.078445
	3	+.24		+ .357620		− 1.368275		
	1	−.09		+ .164203		− 1.080198		
90	2	−.44	−.2500	− .197115	− .002885	− 1.115161	− 3.198063	−.250902
	3	−.22		+ .030027		− 1.002704		
	1	+.24		+ .164203		− 1.080198		
91	2	+.06	+.0800	− .020008	+ .001397	− 1.001200	− 3.142172	+.079555
	3	−.06		− .142798		− 1.060774		
	1	+.05		+ .139303		− 1.057853		
92	2	+.12	−.0867	+ .215139	− .034122	− 1.137879	− 3.632231	−.096094
	3	−.43		− .389164		− 1.436499		
	1	−.10		− .189672		− 1.106718		
93	2	+.03	+.0833	− .053451	+ .007625	− 1.008563	− 3.300383	+.080990
	3	+.32		+ .250748		− 1.185102		
	1	−.08		− .212064		− 1.133055		
94	2	+.26	+.1233	+ .139303	− .005763	− 1.057853	− 3.204354	+.125098
	3	+.19		+ .066998		− 1.013446		
	1	−.40		− .183494		− 1.098764		
95	2	+.30	−.2233	+ .720642	+ .143083	− 2.415765	− 4.961730	−.194463
	3	−.57		− .394067		− 1.447201		
	1	+.01		+ .080515		− 1.019406		
96	2	−.13	−.0700	− .060216	+ .000291	− 1.010865	− 3.031471	−.069904
	3	−.09		− .020008		− 1.001200		
	1	−.04		− .083882		− 1.021059		
97	2	+.28	+.0433	+ .250748	+ .009877	− 1.185102	− 3.279518	+.040288
	3	−.11		− .156989		− 1.073357		
	1	−.21		+ .097612		− 1.028495		
98	2	−.18	−.3067	+ .128767	− .018516	− 1.049476	− 3.240918	−.312413
	3	−.53		− .235018		− 1.162947		
	1	−.02		+ .050125		− 1.007531		
99	2	−.12	−.0700	− .050125	+ .000000	− 1.007531	− 2.015062	−.070000
	3	−.07		+ .000000		− 0.000000		
	1	−.09		− .030027		− 1.032758		
100	2	−.06	−.0600	+ .000000	+ .000000	− 0.000000	− 2.065516	−.060000
	3	−.03		+ .030027		− 1.032758		

In order to analyze the results of the calculations on the preceding ten pages it is necessary that we find the theoretical variance for $\bar{x}$ and $\hat{m}$ from the formulae derived for this purpose ·lier in this paper.

From (7) page 90, we get after setting $a=1$, $p=2$, and $n=3$,

(10) $$\sigma_{\bar{x}}^2 = \frac{1}{21} = .047619$$

and from (8) page 90 after similar substitutions

(11) $$\sigma_{\hat{m}}^2 = \frac{1}{30} = .033333$$

From (9) page 90 we get for the efficiency of the mean when

(12) $$p = 2 \quad E = .70$$

Now by actual calculation from the ungrouped data the mean square deviation from zero which we shall designate as

(13) $$\frac{\Sigma \bar{x}^2}{N} = .048611$$

and

(14) $$\frac{\Sigma m_2^2}{N} = .047612$$

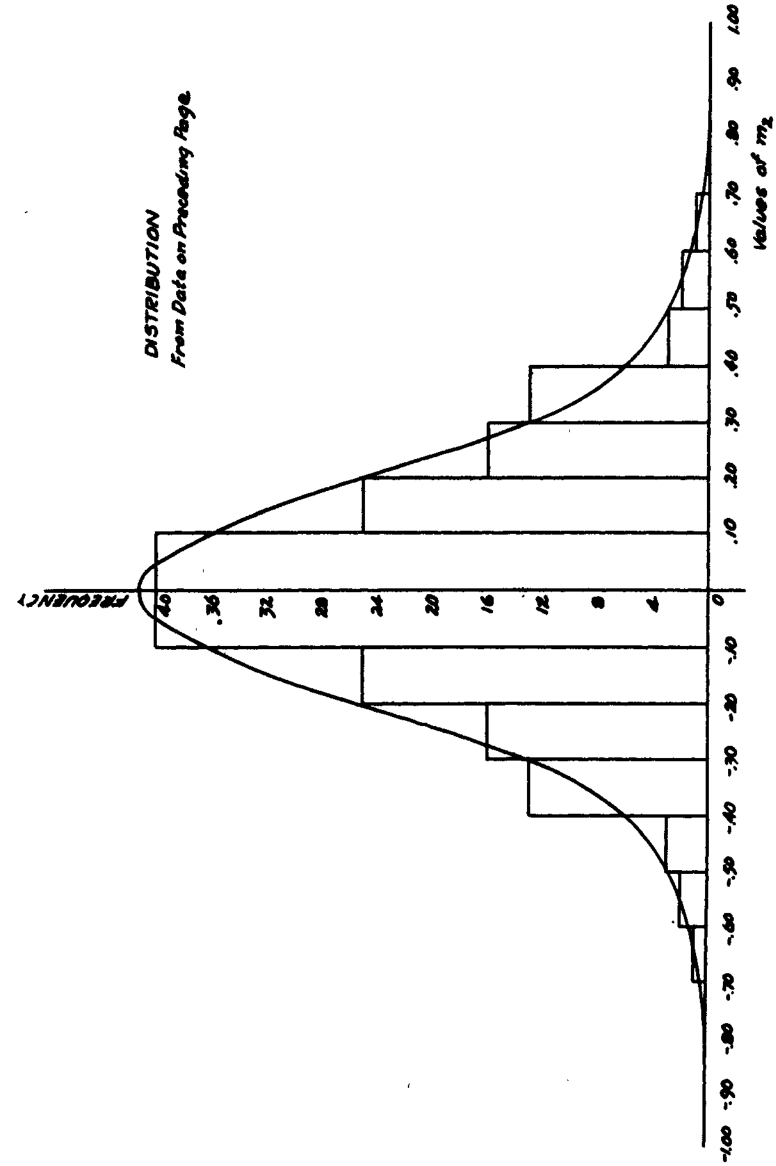
Comparing (10) and (13) it is evident that such a difference can be said to be well within the limits of random sampling. The difference between (11) and (14) is of such magnitude that we can not say that it might be expected in the course of random sampling. Now since m_2 is our estimate of $\hat{m}$ it is evident that either a single approximation by Newton's method is not adequate to give the best results or the approximation,

$$\sigma_{\hat{m}}^2 = \frac{a^2(p-1)}{np(2p+1)},$$

to the variance is not valid in the case of small samples. That the latter seems to be the case the writer firmly believes. The reason for this belief lies in the fact that in a subsequent case a sample of 3 was examined by Newton's method, and starting with $m_1 = \bar{x} = -.07$ the values $m_2 = -.066879$ and $m_3 = -.066791$ were obtained. There is not sufficient improvement here to cause one to suppose that by taking a third approximation we would obtain a variance in keeping with the one derived from theory. Also considering the mean square deviation from zero for $\bar{x}$ and m_2 it seems that the gain in accuracy to be expected from the use of the method of maximum likelihood solution instead of the arithmetic mean is not sufficiently great, in the case of samples of three, to warrant the additional labor involved in calculation. We must be sure, however, that in using the arithmetic mean we are using an approximation to m which complies with the qualifications given on pages 95 and 96. A graph of the distribution as found from the calculations is given on page 106. The histogram represents the grouped data while the smoothed curve is a rough approximation to the actual form of the distribution.

Histogram Data

	m_2	+	-	Totals
	0.00	2	2	0
.01	−.10	16	20	40
.11	.20	8	17	25
.21	.30	4	12	16
.31	.40	4	9	13
.41	.50	1	2	3
.51	.60	1	1	2
.61	.70	0	1	1
.71	.80	0	0	0
.81	.90	0	0	0
.91	1.00	0	0	0
		36	64	100


DISTRIBUTION
From Data on Preceding Page
FREQUENCY
Values of m_2
0 4 8 12 16 20 24 28 32 36 40
-1.00 -.90 -.80 -.70 -.60 -.50 -.40 -.30 -.20 -.10 0 .10 .20 .30 .40 .50 .60 .70 .80 .90 1.00

The totals have been used in the histogram which has been forced to be symmetrical so as to give the effect of a sample of twice the size.

The tabulation of the histogram data draws attention to the great excess of samples having negative values for $\bar{x}$ and m_2. This has caused the writer no little concern. In examining the signs of the observations we note that there are 183 − and only 117 + values. Assuming + and − values to be equally likely this gives a deviation from 150 of 33 or 3.81 times its standard error which is incredible. It seems therefore that Tippett's numbers are not random in this respect, and that it perhaps would have been better to toss a coin to determine the signs.

As a final check and in an effort to place the type of the distribution the value of β_2 has been calculated, and found to be

$$\beta_2 = 3.056$$

β_1 was not calculated as the excess of negative signs would lead to an erroneous value. There is every reason to believe that β_1 should be zero. These facts suggest that the curve is very near to the normal curve, but perhaps slightly more leptokurtic. But, why, if this is the case, there is not better agreement between (11) and (14) page 104 the writer is unable at the present to state.