DISTRIBUTIONS OF CERTAIN FREQUENCY CONSTANTS IN SAMPLES FROM NON-NORMAL POPULATIONS

BY K. C. CHERIYAN
Statistical Laboratory, Calcutta

INTRODUCTION

In this paper a study of the sampling distribution of the correlation coefficient in samples from correlated Pearsonian type III populations is made by means of experimental sampling. Using the transformation $z = \frac{1}{2}\{\log_e (1 + r) - \log_e (1 - r)\}$ to the observed values of r, the sampling distributions of z are obtained. The frequencies and frequency constants of these distributions are then compared with corresponding 'normal theory' values. The sampling distributions of variance and mean are also recorded in order to see how far they agree with known theoretical results.

THE DISTRIBUTION OF THE CORRELATION COEFFICIENT

The historical background to this problem may be briefly described as follows. The theoretical distribution of the correlation coefficient, when samples are drawn from an infinite bivariate normal population, was given by Fisher in 1915. As the mathematical expression defining this distribution was not in a form suitable for the evaluation of frequencies in different class-intervals, 'A Co-operative Study' was made by several authors (1915) to construct tables giving the relative frequencies for different values of r. Later in 1929, E. S. Pearson showed that Fisher's z-transformation could be used to obtain frequencies in the different class-intervals of the r-distribution by entering values of $(z - \text{mean } z)/\sigma_z$ in any table of the normal probability integral. In particular, he took the first and second approximations for the values of mean z and σ^2_z. These two methods have since been used to compare values of r obtained in samples from non-normal populations with normal theory values. Student's work (1908) in this direction is prior to that of Fisher, but the population he used for experimental sampling was only approximately normal. E.S. Pearson (1929 & 1932) has contributed several articles in which it is observed that for parent populations with different values of ρ, there is close agreement between observation and 'normal theory'. The same result was obtained by Rider (1932) who used rectangular and triangular populations. But as Hey (1928) points out, "No attempt appears to have been made to carry out experimental sampling from a bivariate population in which the distribution surface is not normal and in which the correlation is high". Since the theoretical distribution of r has been established only in the case where the parent population is normal, experimental sampling from all other types of populations will be welcome as they help to determine whether tests of significance based on 'normal theory' assumptions are relevant or not.

EXPERIMENTAL WORK

In this paper, bivariate type III populations with varying degrees of skewness and correlation are constructed from the squares of normal deviates.† Since each of the squared values has a distribution given by the law, $y = y_0 x^{-1/2} e^{-x}$, the sum of n of these will be distributed according to the law, $y = y_0 x^{(n-1)/2} e^{-x}$. Thus, if it be required to construct a population in two variables with $\rho = 3/4$, we take groups of five from among the squares of normal deviates, add the first four and the last four and denote them x and y. This device provides

*Part of a thesis approved by the University of Madras for the M. Sc. Degree.

†This table was formed by A. N. Krishnan Nair (1941).

an easy way of constructing populations with any degree of skewness and any desired value for ρ. In another paper (1941) I have obtained mathematical expressions for this type of frequency distributions as also for their moment generating functions. In a recent paper dealing with the same problem Krishnan Nair (1941) has constructed two populations with marginal distributions $y = y_0 e^{-x}$ and $y = y_0 x e^{-x}$. In each case $\rho = 0$. It is observed that in both cases normal theory values provide good fit to sampling distributions of r, and also much better fit than either of the approximations. The problem attempted here is a continuation of this work by sampling from populations in which the variates are highly correlated.

The following three bivariate populations were constructed according to the method outlined above. The betas and skewnesses in the following table are those of the marginal distributions.

	Marginal Distribution	β_1	β_2	Skewness	Correlation
Population I	$y=y_0\, e^{-x}$	4·0	9·0	1·0000	1/2
Population II	$y=y_0\, x\, e^{-x}$	2·0	6·0	0·7070	3/4
Population III	$y=y_0\, x^{1/4}\, e^{-x}$	0·9	4·3	0·4814	8/9

Two hundred samples of five were taken from each of the three populations and the correlation coefficient was calculated for each sample. Thus we get three experimental distributions of r, the total frequency in each distribution being 200. To find how far these distributions show agreement with 'normal theory', we obtain the frequencies for the latter from the tables provided in 'A co-operative study' (13). Frequencies were also obtained by employing the following approximations for the means and variances of transformed r, and using Sheppard's tables of the normal probability integral.

Approximation I.

$$\text{Mean } z = \xi + \frac{\rho}{2(n-1)}, \quad \sigma^2_z = \frac{1}{n-3}$$

Approximation II.

$$\text{Mean } z = \xi + \frac{\rho}{2(n-1)} \left\{1 + \frac{1+\rho^2}{8(n-1)} + \ldots\right\}$$

$$\sigma^2_z = \frac{1}{n-1} \left\{1 + \frac{4-\rho^2}{2(n-1)} + \frac{176-21\rho^2-21\rho^4}{48(n-1)^2} + \ldots\right\}$$

$2(n-1)$ $48(n-1)^2$

Table 1 gives the theoretical and observed distributions of r. Table 2 gives the observed and expected values of the first two moments and betas.

It is seen that in these tables, the χ^2-values indicate very poor agreement between observation and normal theory. But examining the tables more closely, we find that this is mainly due to a few unusually high values among the observed frequencies. This might very well be due to errors of sampling or an inadequate total frequency. In all other class-intervals, the agreement between observation and theory is remarkably good. Further, as we are sampling from populations with high positive correlations, we are only concerned with the distribution of the observed frequencies in the immediate neighbourhood of the theoretical value. Considering the total frequencies at either end of the range, we have the grouped values shown in Table 3. These results are in fairly close agreement between observation and theory, but it should be noted that as the correlation increases, the agreement becomes less satisfactory.

TABLE 1. DISTRIBUTION OF r FOR POPULATIONS I, II AND III

central values of r	Population I ($\rho=0·5$)				Population II ($\rho=0·75$)				Population III ($\rho=8/9$)			
	observed frequencies	expected: normal theory	expected: approximation I	expected: approximation II	observed frequencies	expected: normal theory	expected: approximation I	expected: approximation II	observed frequencies	expected: normal theory	expected: approximation I	expected: approximation II
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
−0·93	1	0·6	0·3	0·1	1	0·1	0·1	0·2	—	0·0	0·0	0·-
−0·85	—	1·1	1·2	0·7	—	0·2	0·3	0·7	—	0·0	0·0	0·1
−0·75	1	1·6	2·1	1·4	—	0·4	0·6	1·1	1	0·1	0·1	0·2
−0·65	3	2·0	2·9	2·2	1	0·5	0·6	1·4	—	0·1	0·1	0·3
−0·55	3	2·6	3·6	2·9	—	0·7	1·6	1·8	—	0·1	0·2	0·5
−0·45	1	3·1	4·3	3·7	—	0·8	1·7	2·1	—	0·2	0·2	0·6
−0·35	6	3·8	4·9	4·5	2	1·0	2·1	2·6	—	0·2	0·4	0·8
−0·25	5	4·6	5·7	5·4	—	1·3	2·7	2·9	—	0·2	0·6	1·0
−0·15	7	5·6	6·5	6·3	2	1·7	3·1	3·4	—	0·3	0·7	1·3
−0·05	7	6·6	7·3	7·3	4	2·3	4·2	3·9	—	0·6	0·9	1·6
+0·05	18	6·8	8·3	8·4	6	3·0	4·8	4·6	1	0·8	1·3	1·9
0·15	3	9·3	9·3	9·8	3	3·4	5·9	5·3	—	1·2	1·7	2·5
0·25	11	11·1	10·8	11·3	6	5·1	7·2	6·3	—	1·7	2·4	3·1
0·35	11	13·7	12·1	13·0	9	6·7	8·9	7·6	3	2·4	3·3	3·8
0·45	11	15·8	13·9	15·1	13	9·6	11·2	9·4	5	3·7	4·8	5·7
0·55	16	18·6	16·2	17·6	12	13·6	14·3	11·9	3	6·9	7·1	7·6
0·65	17	21·8	19·1	20·6	19	19·5	15·3	15·7	11	11·7	11·2	11·1
0·75	14	24·8	22·7	24·1	26	29·2	29·6	21·9	22	20·3	19·4	12·8
0·85	26	26·4	26·7	26·9	29	45·4	37·8	34·3	27	43·8	39·4	24·2
0·93	39	19·6	22·4	18·7	64	52·2	48·0	63·1	127	108·2	104·2	103·9
Total	200	199·3	200·1	200·0	200	196·6	200·0	200·1	200	201·6	199·5	200·0
χ^2		50·02	22·09	44·50		17·21	24·54	11·27		14·24	19·88	23·56
n'		15	15	15		11	11	11		8	8	8
P		0·000	0·077	0·000		0·072	0·006	0·379		0·040	0·006	0·001

TABLE 2. FREQUENCY CONSTANTS WITH CORRESPONDING 'NORMAL THEORY' VALUES

frequency constants	Population I		Population II		Population III	
	observed	expected	observed	expected	observed	expected
Mean	0·4590	0·4317	0·6770	0·7012	0·8580	0·8560
σ	0·4528	0·4239	0·3331	0·3025	0·1547	0·1927
β_1	0·0323	1·0315	3·0027	4·1720	18·3945	12·1821
β_2	2·7574	3·4101	6·6138	7·9582	31·8453	20·4275
Skewness	2·3048	1·7097	1·3761	1·0502	1·8713	2·0356

TABLE 3. GROUPED DISTRIBUTION OF r AT EITHER END OF RANGE

Total frequency		observed	normal theory	approximation I	approximation II
Population I	Below zero	34	31·4	38·8	34·5
	Above 0·7	79	71·0	71·8	69·7
" II	Below zero	10	9·0	17·0	20·0
	Above 0·8	97	97·6	85·6	97·4
" III	Below zero	1	1·9	3·1	6·4
	Above 0·8	154	152·0	143·6	140·1

In table 1, large deviations from expected values are found only in the class-intervals 0·0 to 0·1, 0·1 to 0·2 and 0·9 to 1·0 for Population I. For values of r between 0·2 and 0·7, the deviations from theory are very small. In the same table, for Populations II and III serious discrepancies are found to occur only towards the positive end of the range.

An important result revealed by this study is that 'normal theory' gives a fairly good fit in the immediate neighbourhood of the population correlation only when this correlation is not very high. For Population I ($\rho=0\cdot5$,) for a deviation from the theoretical as great as 0·25, we are justified in using normal theory for tests of significance.

For Population II ($\rho=0\cdot75$), a sufficiently close agreement between observation and 'normal theory' is noticeable only for values of r below 0·75, and above 0·10. For any observed correlation falling outside this range, 'normal theory' tables cannot be used, as the total frequency in these tables is not large enough for a discussion of the frequencies at the 5 and 1 per cent levels. For Population III ($\rho=0\cdot88$), 'normal theory' values show agreement only between 0·3 and 0·8. The observed agreement is not sufficient to justify the use of 'normal theory' assumptions.

The constants of the sampling distributions indicate very good agreement for means and standard deviations. The values of the betas are also fairly close, except for Population III. According to normal theory, the skewness of the r-distribution increases as ρ increases. But the observed skewness is maximum for the first distribution and minimum for the second.

DISTRIBUTION OF z.

The theoretical distribution of z has been shown to be approximately normal in the case where the parent population is normal. But if this property is found to persist when the parent universe is a skew one, then we can safely apply tests of significance to values of transformed r. Table 4 gives distributions of z where the z-values were obtained by entering Fisher's table (4) with the corresponding value of r.

TABLE 4. DISTRIBUTION OF z

Central values of z	Population I	Population II	Population III observed	Population III expected	Central values of z	Population I	Population II	Population III observed	Population III expected
Below −1·40	—	1	1	—	Below 1·70	10	12	21	21·0
−1·30	1	—	—	—	1·90	1	11	23	21·1
−1·10	1	—	—	—	2·10	—	9	19	16·9
−0·90	2	—	—	—	2·30	—	4	9	12·2
−0·70	2	1	—	0·1	2·50	—	5	9	7·9
−0·50	4	1	—	0·2	2·70	—	8	2	4·9
−0·30	9	2	—	0·4	2·90	—	2	5	2·5
−0·10	12	6	—	1·0	3·10	—	4	3	1·3
0·10	24	8	1	2·0	3·30	—	1	3	0·6
0·30	20	13	3	3·9	3·50	—	—	3	0·2
0·50	20	19	6	6·8	3·70	—	—	—	0·1
0·70	23	16	7	10·7	3·90	—	—	1	—
0·90	27	31	20	15·3	4·10	—	—	3	—
1·10	20	14	13	19·8	4·30	—	—	3	—
1·30	12	10	19	23·3	4·70	—	—	1	—
1·50	12	13	25	25·0	Total	200	200	200	200·2

$\chi^2=30\cdot40$, $n'=13$, $P=0\cdot001$

It is seen that the betas of the first two distributions agree well with theory. For the third distribution we observe deviations for the variance and betas. The expected

TABLE 5. PARAMETERS OF THE DISTRIBUTION OF z.

Parameter	Population I		Population II		Population III	
	observed	expected	observed	expected	observed	expected
Mean	0·8190	0·8143	1·1690	1·0713	1·7310	1·5337
μ_2	0·3920	0·4213	0·6925	0·4100	0·7593	0·4064
μ_3	0·1616	0·0024	0·1652	0·0000	0·4727	0·0032
μ_4	0·4530	0·5600	1·5191	0·5374	2·7799	0·5178
β_1	0·4333	0·0003	0·0655	0·0000	0·5100	0·0006
β_2	2·9479	3·7945	3·1861	3·0177	4·8192	3·0325
Skewness	0·6234	0·0059	0·1211	0·0000	0·2320	0·0069

values in this case are those of the corresponding z-distributions assuming the parent universe to be normal. They were obtained from normal probability tables by entering values of $(z - \text{mean } z)/\sigma_z$ where the means and the standard deviations are the theoretical values. We observe that agreement is good everywhere except for values of z above 2·60. The total observed frequencies on either side of the theoretical mean are fairly equally balanced for the three distributions as is shown below

Frequency	Population I	Population II	Population III
Below mean	95	98	95
Above mean	105	102	105

In the last case we also find that agreement is good in the immediate neighbourhood of the theoretical z-values. For correlations obtained from highly correlated skew populations, the z-transformation might provide a better test of significance than the corresponding 'normal theory' tables when the observed z-value does not deviate greatly from the theoretical.

DISTRIBUTION OF VARIANCE

The exact distribution of variance in samples from non-normal populations is not known. But as Le Roux (1931) has observed, it may be supposed to conform to either type III, type V or type VI of Pearson's curves. Tchouproff (1916) first gave the formulae for moments of the variance distribution and later Church (1925) derived them again following 'Student's' shorter and more direct method. For samples of size 5, the values of the first two moments and betas are as given below.

$$M'_1 = \mu_2 \frac{4}{5} \cdot M_2 = \mu_2 \left(\frac{4}{5}\right)^2 (\beta_2 - \tfrac{1}{2})$$

$$B_1 = \frac{\left(\beta_4 - \frac{25\beta_2}{4} - \frac{1}{2}\right)^2}{5(\beta_2 - \frac{1}{2})^3}$$

$$B_2 = \frac{\left(\beta_6 + 2\beta_4 - 20\beta_3 + \frac{\beta^2_2}{8} - \frac{195\beta_2}{4} + \frac{145\beta_1}{2} + \frac{420}{8}\right)}{5(\beta_2 - \frac{1}{2})^2}$$

Where M'_1, M_2, B_1, B_2, are the moments and betas of the variance-distribution, and μ_2, β_1, β_2, ..., β_6 are parameters of the parent population. The observed distributions of σ^2_x and σ^2_y from each of the three populations as well as the observed and expected values of constants of the variance-distribution are given in the following tables 6 and 7.

TABLE 6. FREQUENCY DISTRIBUTION OF VARIANCES OF x AND y.

Central Values	Population I		Central Values	Population II		Central Values	Population III	
	σ_x^2	σ_y^2		σ_x^2	σ_y^2		σ_x^2	σ_y^2
0·25	24	20	0·5	18	18	2·0	16	12
0·75	32	23	1·5	20	22	4·0	18	23
1·25	31	34	2·5	26	28	6·0	28	24
1·75	12	14	3·5	25	29	8·0	23	21
2·25	15	26	4·5	24	20	10·0	16	22
2·75	17	13	5·5	16	8	12·0	17	17
3·25	8	15	6·5	16	16	14·0	12	15
3·75	11	8	7·5	10	14	16·0	13	11
4·25	8	6	8·5	6	11	18·0	10	7
4·75	2	1	9·5	11	8	20·0	11	12
5·25	5	6	10·5	4	4	22·0	6	6
5·75	4	4	11·5	4	2	24·0	3	8
6·25	6	1	12·5	2	4	26·0	7	2
6·75	3	2	13·5	4	2	28·0	3	8
7·25	5	3	14·5	4	5	30·0	3	2
7·75	1	1	15·5	—	3	32·0	4	4
8·25	3	—	16·5	1	—	34·0	4	1
8·75	5	3	17·5	2	3	36·0	—	—
9·25	1	1	18·5	—	4	38·0	1	2
9·75	1	4	19·5	—	—	40·0	—	2
10·25	—	—	20·5	1	2	42·0	1	—
10·75	1	1	21·5	1	—	44·0	—	1
11·25	—	—	22·5	1	—	46·0	1	1
11·75	2	—	23·5	—	—	above 47·0	4	8
12·25	1	1	24·5	1	—			
12·75	1	2	25·5	—	1			
above 13	1	5	above 26	4	3			
Total	200	200	Total	200	200	Total	200	200

From table 7 we note that the observed B_1 and B_2 for Population I is considerably less than what might be expected. This might possibly be due to the excessive skewness in Population I. Another point of difference is that the observed standard deviations in the third case are both less than half the expected value. All the other values agree fairly well.

TABLE 7. CONSTANTS IN THE DISTRIBUTION OF THE VARIANCE.

Constants	Population I			Population II			Population III		
	observed		expected	observed		expected	observed		expected
	x	y		x	y		x	y	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mean	2·0650	2·4000	3·2000	7·1300	6·3250	6·4000	11·1450	10·9400	14·4000
$\sqrt{M_2}$	2·8437	4·1576	4·1721	6·0844	6·1400	6·7124	8·1130	5·6133	13·6063
B_1	2·3240	1·3058	18·3696	7·9944	12·0818	10·6459	6·7241	3·4766	6·4915
B_2	5·0631	4·6468	35·0585	13·6248	22·3303	20·1694	12·4583	7·2799	10·4273

DISTRIBUTION OF MEAN

The theoretical distribution of the mean when samples of size n are drawn from a type III population of the form $y=\dfrac{x^{p-1}e^{-x}}{\Gamma(p)}$ was proved by Irwin to be (1927) $y=\dfrac{ne^{-n\bar{x}}\,(n\bar{x})^{np-1}}{\Gamma(np)}$

Since the marginal distributions of the parent populations used in this paper are $y=y_0\,e^{-x}$, $y=y_0\,x\,e^{-x}$ and $y=y_0\,x^{7/2}\,e^{-x}$, the corresponding mean-distributions will be given by

$$y=\{(5u)^4\,e^{-5u}\,d(5u)\}/\Gamma(5) \qquad .. \quad (1)$$

$$y=\{(5u)^9\,e^{-5u}\,d(5u)\}/\Gamma(10) \qquad .. \quad (2)$$

$$y=\{(5u)^{21\cdot5}\,e^{-5u}\,d(5u)\}/\Gamma(22\cdot5) \qquad .. \quad (3)$$

In calculating the expected moments from these equations it has to be remembered that the unit used in the parent distribution is really $\chi^2/2$ and so for comparison with observed values, the moment μ_r obtained from this should be multiplied by 2^r for all values of r.

Tables 8 and 9 give the observed distributions of $\bar{x}$ and $\bar{y}$ from the three populations as well as the observed and expected values of the constants.

TABLE 8. FREQUENCY DISTRIBUTION OF THE MEANS.

central values	Population I $\bar{x}$	Population I $\bar{y}$	central values	Population II $\bar{x}$	Population II $\bar{y}$	central values	Population III $\bar{x}$	Population III $\bar{y}$
0·5	4	5	1·25	2	2	4·5	1	2
0·7	12	7	1·75	6	6	5·5	7	3
0·9	7	11	2·25	14	11	6·5	13	21
1·1	14	14	2·75	15	20	7·5	36	40
1·3	12	13	3·25	36	31	8·5	42	39
1·5	29	26	3·75	27	35	9·5	42	32
1·7	15	15	4·25	25	34	10·5	20	29
1·9	23	15	4·75	18	26	11·5	17	19
2·1	8	16	5·25	21	14	12·5	6	5
2·3	23	20	5·75	6	11	13·5	4	8
2·5	[illegible]	17	6·25	2	1	14·5	3	1
2·7	14	7	6·75	5	4	15·5	—	1
2·9	6	14	7·25	1	3	16·5	—	—
3·1	7	9	7·75	2	1	17·5	1	—
3·3	6	1	8·25	—	1			
3·5	4	3	8·75	—	—			
3·7	1	2						
3·9	3	1						
4·1	—	—						
4·3	1	1						
4·5	—	1						
above 4·5	1	2						
Total	200	200	Total	200	200	Total	200	200

Agreement of the first two moments with theory is everywhere perfectly good. The higher moments show deviations only in the case of the distributions of $\bar{x}$, but no explanation for this behaviour is evident from an inspection of the frequencies. Observed β_1 and β_2 give fairly good approximations to theoretical values in all the three cases.

TABLE 9. CONSTANTS IN THE DISTRIBUTION OF THE MEAN.

constants	Population I observed x	Population I observed y	Population I expected	Population II observed x	Population II observed y	Population II expected	Population III observed x	Population III observed y	Population III expected
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Mean	1·0500	1·7420	2·0000	3·5375	3·4075	4·0000	9·1250	9·1050	9·0000
μ_2	0·6086	0·6392	0·8000	1·3559	1·4867	1·6000	3·8211	3·8657	3·6000
μ_3	0·3506	0·5277	0·6400	0·6706	1·0051	1·2800	5·2403	3·7697	2·8800
μ_4	1·4401	2·1101	2·6880	4·5058	8·2453	9·2160	62·7080	45·7804	42·3360
β_1	0·4112	1·0062	0·8000	0·1804	0·3074	0·4000	0·4022	0·2422	0·1778
β_2	3·2217	5·1662	4·2000	2·4508	3·7304	3·6000	4·2989	3·0325	3·2667

SUMMARY AND CONCLUSIONS

In this paper an attempt is made to study by means of experimental sampling the distribution of r when the parent population is of the type III form and where the variates are highly correlated. For this purpose, 200 samples of five were taken from each of three different populations whose distribution laws are known, and having correlations 0·5, 0·75 and 0·89. The distributions of transformed r as well as those of variance and mean have also been recorded to see how far their first few moments and betas agree with theoretical results.

It has to be noted that skewness is maximum for Population I and minimum for Population III. To study the significant changes in the r-distribution due to changes in population skewness or population correlation, we should have taken either all populations with the same correlation and varying skewness or vice versa. Further it might be of interest to draw samples from populations with different marginal distributions. Such detailed analyses are, however, beyond the scope of this paper. It might be mentioned however that in a previous investigation of this nature (1941), the change in population skewness is found to have no apparent effect on the symmetry of the r-distribution when the population correlation is kept constant at zero. But it is observed here that just as in the normal case, the r-distruibution becomes more and more asymmetrical as ρ increases from 0 to 1.

Some of the main results emerging from this study can be stated as follows:—

(1) Close agreement with 'normal theory' frequencies is found only in the ranges 0·2 to 0·7, 0·1 to 0·8 and 0·3 to 0·8 respectively for the three distributions of r. At either end of the range, the total frequencies agree well with theory. For Population III, the frequencies at the lower end of the range do not show agreement either with 'normal theory' or the approximations.

(2) The first two sampling distributions of z show good agreement with 'normal theory' as is evident from the values of the frequency constants. In the third case, the only divergence from theory is for frequencies above 2·60. The observed distribution is slightly more leptokurtic than would be expected from 'normal theory'. In all three cases, the total frequencies on either side of the theoretical means are fairly equally balanced.

(3). Among the variance-distributions, good agreement with theory is noticed only for sampling distributions from Population II.

(4) All the mean-distributions show good agreement with theoretical values.

for sampling distributions from Population II.

(4) All the mean-distributions show good agreement with theoretical values.

REFERENCES.

1. CHERIYAN K. C. (1941) A bivariate correlated gamma-type distribution function. *Journal of the Indian Mathematical Society.* (New Series) 5 (3). 133-144.

2. CHURCH, A. E. R. (1925) On the moments of the distribution of squared standard deviations for small samples of N drawn from an indefinitely large population. *Biometrika* 17, 79.

3. FISHER, R. A. (1915) The frequency distribution of the values of the correlation coefficient in samples from an indefinitely large population. *Biometrika* 10, 507-521.

4. FISHER, R. A. (1921) On the probable error of a coefficient of correlation deduced from small samples. *Metron* 1 (4) 3-32.

5. REY, G. B. (1928) A new method in experimental sampling illustrated on certain non-normal populations. *Biometrika* 30, 83-80.

6. IRWIN, J. O. (1927) On the frequency distribution of means of samples from a population having any law of frequency with finite moments. *Biometrika* 19, 228.

7. KRISHNAN NAIR, A. N. (1941) Distribution of "Student's" 't' and the correlation coefficient in samples from non-normal populations *Sankhyā* 5 (4), 383-400.

8. LE ROUX, J. M. (1931) A study of the distribution of the variance in small samples. *Biometrika* 23, 134-190.

9. LEONE CHESIRE, ELENA OLDIS AND PEARSON, E. S. (1932) Further experiments on the sampling distribution of the correlation coefficient. *Journal of the American Statistical Association* 27, 121-128.

10. PEARSON, E. S. (1929) Some notes on sampling tests with two variables. *Biometrika* 21, 337-360.

11. PEARSON, E. S. (1932) The test of significance for the correlation coefficient: Some further results. *Journal of the American Statistical Association* 27, 424-426.

12. RIDER, P. R (1932) On the distribution of the correlation coefficient in small samples. *Biometrika* 24, 382-403.

13. SOPER, H. E. & OTHERS, (1913) On the distribution of the correlation coefficient in small samples. Appendix II. to the papers of 'Student' and R. A. Fisher—A co-operative study. *Biometrika* 11, 328-413. Tables 379-413.

14. STUDENT (1908) On the probable error of a correlation coefficient. *Biometrika*, 6 302.

15. TCHOUPROFF (1916) On the mathematical expectation of the moments of frequency distributions. *Biometrika* 12, 103-104.

Paper received: 16 September, 1942