MISCELLANEA

RELATIVE EFFICIENCY OF GAUGING AND EXACT MEASUREMENT IN ESTIMATING THE PROPORTION OF A POPULATION BETWEEN GIVEN LIMITS

By DES RAJ
Indian Statistical Institute, Calcutta

1. Introduction

In order to control the quality of a manufactured product, we may either actually measure the articles or simply gauge them. That it is quicker, easier and therefore cheaper to gauge an article than to measure it is obvious. What is not so widely understood is that the efficiency of the technique of gauging for controlling the quality of a product is also generally high. Stevens (1948) for example, shows that it is possible to achieve with ten gauged articles the same sensitivity of control of the mean as is given by eight articles exactly measured. The object of this note is to study the allied problem of estimating the proportion of the product lying between two specified limits. The relative efficiency of gauging as compared with exact measurement has been obtained for a wide class of distributions.

Let there be a population

$$f(x, \theta) \qquad \dots \quad (1)$$

containing the parameters $\theta = (\theta_1, \theta_2, \dots, \theta_k)$.

The object is to estimate

$$P = \int_a^b f(x, \theta)dx, \qquad \dots \quad (2)$$

where a and b are two specified constants.

Suppose we have a pair of gauges set to the values a and b of the variate x. Out of a sample of size N drawn from (1), let n fall between a and b. The estimate of P based on the gauging method is

$$p_{gaug.} = \frac{n}{N} \qquad \dots \quad (3)$$

and the variance of the estimate is

$$P(1-P)/N. \qquad \dots \quad (4)$$

If, instead of gauging the articles we actually measure them and estimate the parameters θ by θ^*, the estimated proportion would be

$$p_{meas.} = \int_a^b f(x, \theta^*)dx. \qquad \dots \quad (5)$$

Using the formula (Rao, 1952), large sample variance of (5), under usual assumptions, would be given by

$$\sum\sum\left\{\left[\int_a^b \frac{\partial f}{\partial \theta_i^*}\, dx\right]\left[\int_a^b \frac{\partial f}{\partial \theta_j^*}\, dx\right]\right\}_{\theta^*=\theta} \text{Cov}\,(\theta_i^*, \theta_j^*). \qquad \text{... (6)}$$

The relative efficiency of the gauging method is then the ratio of (6) and (4).

2. Efficiency of gauging for the type III population

The type III population is given by

$$f(x)dx, \quad m-\frac{2\sigma}{\alpha_3} < x < \infty,\ 0 < \alpha_3 < 2 \qquad \text{... (7)}$$

where

$$f(x) = \frac{C}{\sigma}\left\{1+\frac{\alpha_3}{2}\,\frac{x-m}{\sigma}\right\}^{\frac{4}{\alpha_3^2}-1} e^{-\frac{2}{\alpha_3}\frac{x-m}{\sigma}} \qquad \text{... (8)}$$

and

$$C = (4/\alpha_3^2)^{\frac{4}{\alpha_3^2}-\frac{1}{2}}\, e^{-4/\alpha_3^2}\left[\Gamma(4/\alpha_3^2)\right]^{-1} \qquad \text{... (9)}$$

The parameters m, σ and α_3 in (7) are the mean, variance and third standard moment respectively. We are to estimate

$$P = \int_{m+\lambda\sigma}^{m+\mu\sigma} f(x)dx = \int_{\lambda}^{\mu} f(t)dt \qquad \text{... (10)}$$

where

$$f(t) = C\left(1+\frac{\alpha_3}{2}t\right)^{\frac{4}{\alpha_3^2}-1} e^{-2t/\alpha_3}, \quad -\frac{2}{\alpha_3} < t < \infty \qquad \text{... (11)}$$

so that $f(t)$ is the standardised type III curve tabulated by Salvosa (1930). In this case we take

$$p_{\text{meas.}} = \int_{m+\lambda\sigma}^{m+\mu\sigma} \frac{C}{s}\left\{1+\frac{\alpha_3}{2}\left(\frac{x-\bar{x}}{s}\right)\right\}^{\frac{4}{\alpha_3^2}-1} e^{-\frac{2}{\alpha_3}\frac{x-\bar{x}}{s}}\, dx \qquad \text{... (12)}$$

where $\bar{x}$ and s^2 are the sample mean and variance respectively.

We have $\left(\dfrac{\partial}{\partial \bar{x}}\, p_{\text{meas.}}\right)_\theta = \dfrac{l_\lambda - l_\mu}{\sigma}$, $\left(\dfrac{\partial}{\partial s^2}\, p_{\text{meas.}}\right)_\theta = \dfrac{\lambda l_\lambda - \mu l_\mu}{2\sigma^2}$

where l_μ and l_λ are the ordinates of (11) at μ and λ respectively.

Also

$$V(\bar{x}) = \frac{\sigma^2}{N}$$

and in large samples

$$V(s^2) = \frac{\mu_4-\mu_2^2}{N} = \frac{\sigma^4(2+\frac{3}{2}\alpha_3^2)}{N},$$

$$\text{Cov}\,(\bar{x}, s^2) = \frac{\mu_3}{N} = \frac{\sigma^3\alpha_3}{N}.$$

Substituting in (6) we see that (12) is asymptotically normal with mean P and variance given by

$$\frac{1}{N}\Big[\,(l_\mu - l_\lambda)^2 + \Big(\frac{1}{2} + \frac{3}{8}\alpha_3^2\Big)(\mu l_\mu - \lambda l_\lambda)^2 + \alpha_3 (l_\mu - l_\lambda)(\mu l_\mu - \lambda l_\lambda)\,\Big]. \qquad \ldots \quad (13)$$

Putting $\int_{-2/\alpha_3}^{x} f(t)dt = a_x$, we see that the relative efficiency of gauging as compared with exact measurement is

$$\frac{(l_\mu - l_\lambda)^2 + (\frac{1}{2} + \frac{3}{8}\alpha_3^2)(\mu l_\mu - \lambda l_\lambda)^2 + \alpha_3 (l_\mu - l_\lambda)(\mu l_\mu - \lambda l_\lambda)}{(a_\mu - a_\lambda)(1 - a_\mu + a_\lambda)}. \qquad \ldots \quad (14)$$

We shall study this efficiency for different degrees of skewness of this population. For $\alpha_3 = 0$, when the population is normal, the expression for the relative efficiency reduces to $\{(l_\mu - l_\lambda)^2 + \frac{1}{2}(\mu l_\mu - \lambda l_\lambda)^2\}/P(1-P)$, as obtained by Baker (1949). The following cases will be considered :

Case (i) $\mu = \infty$. The relative efficiency in this case is

$$l_\lambda^2[1 + \alpha_3\lambda + (\tfrac{1}{2} + \tfrac{3}{8}\alpha_3^2)\lambda^2][a_\lambda(1 - a_\lambda)]^{-1}. \qquad \ldots \quad (15)$$

The following table gives this efficiency for different degrees of skewness.

TABLE 1. RELATIVE EFFICIENCY OF SINGLE GAUGING FOR TYPE III POPULATION

α_3	0.0		0.2		0.4		0.8	
	P	eff.	P	eff.	P	eff.	P	eff.
2.00	.0228	.393	.0279	.411	.0324	.434	.0396	.493
1.50	.0668	.572	.0716	.573	.0754	.585	.0308	.624
1.00	.1587	.658	.1583	.662	.1572	.671	.1536	.698
0.50	.3085	.654	.2998	.662	.2010	.671	.2736	.688
0.00	.5000	.637	.4867	.630	.4734	.634	.4468	.627
−0.50	.6915	.654	.6826	.645	.6738	.638	.6561	.631
−1.00	.8413	.658	.8418	.663	.8432	.681	.8504	.796
−1.50	.9332	.572	.9393	.586	.9466	.625	.9676	.868
−2.00	.9773	.393	.9820	.384	.9888	.381	.9988	.307

Case (ii) $\mu = -\lambda$. The relative efficiency now is

$$\frac{(l_{-\lambda} - l_\lambda)^2 + (\frac{1}{2} + \frac{3}{8}\alpha_3^2)\lambda^2(l_{-\lambda} + l_\lambda)^2 - \lambda\alpha_3(l_{-\lambda} - l_\lambda)(l_{-\lambda} + l_\lambda)}{(a_{-\lambda} - a_\lambda)(1 - a_{-\lambda} + a_\lambda)}. \qquad \ldots \quad (16)$$

The following table gives this efficiency for different degrees of skewness.

TABLE 2. RELATIVE EFFICIENCY OF SYMMETRICAL GAUGING FOR TYPE III POPULATION

α_3	0.0		0.2		0.4		0.8	
λ	P	eff.	P	eff.	P	eff.	P	eff.
−0.10	.0797	.043	.0796	.043	.0794	.045	.0786	.049
−0.50	.3829	.202	.3829	.208	.3828	.276	.3824	.319
−1.00	.6827	.541	.6835	.552	.6860	.588	.6968	.756
−1.50	.8664	.652	.8670	.669	.8715	.721	.8873	.931
−2.00	.9545	.637	.9550	.547	.9565	.571	.9592	.561

3. EFFICIENCY OF GAUGING FOR THE CAUCHY POPULATION

The Cauchy population is

$$\frac{1}{\pi}\frac{1}{1+(x-\theta)^2}dx, \quad -\infty < x < \infty. \qquad \dots (17)$$

We are required to estimate

$$P = \frac{1}{\pi}\int_{\theta+\lambda}^{\theta+\mu}\frac{dx}{1+(x-\theta)^2} = \frac{1}{\pi}\Big[\tan^{-1}\mu-\tan^{-1}\lambda\Big]. \qquad \dots (18)$$

We consider the estimate $$p_{meas.} = \frac{1}{\pi}\int_{\hat\theta+\lambda}^{\hat\theta+\mu}\frac{dx}{1+(x-\hat\theta)^2} \qquad \dots (19)$$

where $\hat\theta$ is the maximum likelihood estimate of θ.

We have $$\left(\frac{\partial}{\partial\hat\theta}p_{meas.}\right)_\theta = \frac{\mu^2-\lambda^2}{\pi(1+\lambda^2)(1+\mu^2)}$$

and in large samples $$V(\hat\theta) = \frac{2}{N}.$$

Substituting in (6) we see that the large sample variance of (19) is

$$\frac{2}{N\pi^2}\frac{(\mu+\lambda)^2(\mu-\lambda)^2}{(1+\lambda^2)^2(1+\mu^2)^2} \qquad \dots (20)$$

so that the relative efficiency of gauging as compared with exact measurement is given by

$$\frac{2}{\pi^2}\frac{(\mu+\lambda)^2(\mu-\lambda)^2}{(1+\lambda^2)^2(1+\mu^2)^2}\cdot\frac{1}{P(1-P)} \qquad \dots (21)$$

where P is determined from (18).

The following cases will be considered.

Case (i) $\mu = \infty$. The relative efficiency in this case is

$$\frac{2}{\pi^2}\frac{1}{(1+\lambda^2)^2}\cdot\frac{1}{P(1-P)} \qquad \dots (22)$$

where $$P = \frac{1}{\pi}\cot^{-1}\lambda. \qquad \dots (23)$$

The following table gives the relative efficiency for different values of λ.

TABLE 3. RELATIVE EFFICIENCY OF SINGLE GAUGING FOR CAUCHY POPULATION

λ	P	eff.
2.00	.1476	.064
1.50	.1872	.126
1.00	.2500	.270
.50	.3524	.568
.00	.5000	.811

Case (ii) $\lambda = -\mu$. In this case the large sample variance of (19) is zero, which implies that the variance is of a smaller order than n^{-1}. The relative efficiency of gauging as compared with exact measurement is very small.

4. Efficiency of gauging for the normal population

Finally, we shall give detailed tables for the relative efficiency of gauging as compared to exact measurement for the most common distribution, the normal distribution. From these tables, it is found that it is possible to achieve with ten gauged articles the same efficiency for estimating the proportion P as given by six articles exactly measured.

TABLE 4. RELATIVE EFFICIENCY OF SYMMETRICAL GAUGING FOR NORMAL POPULATION

$-\lambda$	P	eff.	$-\lambda$	P	eff.	$-\lambda$	P	eff.
.05	.0399	.021	1.05	.7063	.562	2.05	.9596	.517
.10	.0797	.043	1.10	.7287	.581	2.10	.9643	.495
.15	.1192	.067	1.15	.7499	.598	2.15	.9684	.473
.20	.1585	.092	1.20	.7699	.613	2.20	.9722	.451
.25	.1974	.118	1.25	.7887	.626	2.25	.9756	.428
.30	.2358	.145	1.30	.8064	.636	2.30	.9786	.404
.35	.2737	.174	1.35	.8230	.644	2.35	.9812	.381
.40	.3108	.203	1.40	.8385	.649	2.40	.9836	.358
.45	.3473	.232	1.45	.8529	.652	2.45	.9857	.335
.50	.3829	.262	1.50	.8664	.652	2.50	.9876	.313
.55	.4177	.293	1.55	.8788	.650	2.55	.9892	.291
.60	.4515	.323	1.60	.8904	.645	2.60	.9907	.270
.65	.4843	.353	1.65	.9011	.639	2.65	.9920	.250
.70	.5161	.383	1.70	.9109	.630	2.70	.9931	.230
.75	.5467	.412	1.75	.9199	.619	2.75	.9940	.211
.80	.5763	.440	1.80	.9281	.606	2.80	.9949	.193
.85	.6047	.467	1.85	.9357	.591	2.85	.9956	.176
.90	.6319	.493	1.90	.9426	.574	2.90	.9963	.160
.95	.6579	.518	1.95	.9488	.556	2.95	.9968	.145
1.00	.6827	.541	2.00	.9545	.537	3.00	.9973	.131

TABLE 5. RELATIVE EFFICIENCY OF SINGLE GAUGING FOR NORMAL POPULATION

λ	P	eff.	λ	P	eff.	λ	P	eff.
.00	.5000	.637						
.05	.4801	.637	1.05	.1469	.654	2.05	.0202	.373
.10	.4602	.637	1.10	.1357	.650	2.10	.0179	.353
.15	.4404	.639	1.15	.1251	.644	2.15	.0158	.334
.20	.4207	.640	1.20	.1151	.637	2.20	.0139	.314
.25	.4013	.642	1.25	.1056	.629	2.25	.0122	.295
.30	.3821	.644	1.30	.0968	.620	2.30	.0107	.276
.35	.3632	.646	1.35	.0885	.609	2.35	.0094	.257
.40	.3446	.649	1.40	.0808	.598	2.40	.0082	.239
.45	.3264	.651	1.45	.0735	.585	2.45	.0071	.222
.50	.3085	.654	1.50	.0668	.572	2.50	.0062	.205
.55	.2912	.656	1.55	.0606	.557	2.55	.0054	.189
.60	.2743	.658	1.60	.0548	.542	2.60	.0047	.174
.65	.2578	.660	1.65	.0495	.525	2.65	.0040	.160
.70	.2420	.662	1.70	.0446	.508	2.70	.0035	.146
.75	.2266	.663	1.75	.0401	.490	2.75	.0030	.133
.80	.2119	.663	1.80	.0359	.471	2.80	.0026	.121
.85	.1977	.663	1.85	.0322	.452	2.85	.0022	.110
.90	.1841	.662	1.90	.0287	.433	2.90	.0019	.099
.95	.1711	.661	1.95	.0256	.413	2.95	.0016	.089
1.00	.1587	.658	2.00	.0228	.393	3.00	.0013	.080

It may be noted that for symmetrical gauging the relative efficiency is maximum (65%) when λ is near about 1.50 and P is nearly 0.86, while for single gauging the relative efficiency is maximum (66%) when λ is in the neighbourhood of 0.80 when P is nearly 0.21.

5. Relative efficiency of gauging for equivalent costs

So far we have compared the relative efficiencies of the two methods of estimation assuming that the cost of gauging an article is the same as the cost of actually measuring it. Evidently, this is an over-simplification of the problem. In actual practice the cost of gauging an article would be much less than the cost of actually measuring it. Let

c_1 = cost of gauging an article,

c_2 = cost of measuring an article,

C = total cost.

Then we can gauge n_1 articles or measure n_2 of them for the same cost C, where $n_1c_1 = n_2c_2$. Then the relative efficiency of gauging, as obtained before, would be multiplied by c_2/c_1 throughout. As an illustration, the table below gives the efficiencies for different values of c_2/c_1. In the part enclosed, exact measurement is less efficient than gauging and its relative efficiency is shown.

TABLE 6. RELATIVE EFFICIENCY OF SINGLE GAUGING FOR NORMAL POPULATION FOR EQUIVALENT COSTS

λ \ c_2/c_1	1.5	2.0	2.5	3.0	4.0	5.0	10.0
.10	.956	.785	.628	.523	.392	.314	.157
.50	.981	.765	.612	.510	.382	.306	.153
1.00	.987	.760	.608	.507	.380	.304	.152
1.50	.858	.874	.699	.583	.437	.350	.175
2.00	.590	.786	.983	.848	.636	.509	.254
2.50	.308	.410	.513	.615	.820	.976	.488
3.00	.120	.160	.200	.240	.320	.400	.800

I am grateful to Dr. C. R. Rao under whose guidance this note was written.

References

Baker, G. A. (1949): The variance of the proportion of samples falling within a fixed interval for a normal population. *Ann. Math. Stat.*, 20, 125-126.

Rao, C. R. (1952): *Advanced Statistical Methods in Biometric Research*, John Wiley & Sons, New York, p. 207.

Salvosa, L. R. (1930): Tables of Pearsons' type III function. *Ann. Math. Stat.*, 1, appended.

Stevens, W. L. (1948): Control by gauging. *J. Roy. Stat. Soc.*, Series B, 10, 54-98.

Paper received : August, 1953.

SCALING PROCEDURES IN SCHOLASTIC AND VOCATIONAL TESTS*

By P. K. BOSE and S. B. CHOUDHURY

University of Calcutta

1. Introduction

For testing scholastic and vocational aptitude of different individuals, several kinds of tests are in existence, the commonest one being the ordinary examination system, where the different candidates are subject to written and oral examinations. According to the present system of examination candidates are examined in more than one subject generally; the scores in different subjects are added together and the candidates are ranked or scaled according to the total scores.

A critical analysis of this method of combining the scores will bring out its great drawback. The method does not take into account the differences between the distributions of scores in various subjects. Due to various reasons e.g. (a) intrinsic differences between the subjects, (b) differences in aptitude of candidates in different subjects, (c) differences in standards of examination in different subjects, (d) random fluctuations etc., the distributions of scores in different subjects are usually dissimilar. But the ordinary method of scaling according to total scores assumes that the distributions are identical, which is hardly a case in practice.

A more appropriate method of scaling which may be suggested is to find out the equivalent scores in different subjects with respect to a standard one. Converting scores in different subjects into equivalent scores in one standard subject, one may add up these equivalent scores (instead of the actual scores) to get a total score which may be used for the purpose of scaling.

The problem of scaling can thus be regarded as solved if a method can be evolved to find out the equivalent scores. In this paper an attempt has been made to give a general procedure for finding out the equivalent scores, and to apply this method for solving a practical problem. The method of calculating equivalent scores in some particular situations have been studied by Mahalanobis and Chakravarti (1934), Hussain (1941), and Greenall (1949).

2. Method

In setting up an equivalence between two scores, the following points should be observed : (1) the equivalence set-up should be mutual, and (ii) the set-up should be independent of scale. Bearing these two points in mind we may define equivalent scores as follows:

Two scores in two different subjects will be said to be equivalent when the percentile ranks of these two scores are identical. Let us consider two subjects X and Y, the scores being denoted by x and y respectively. Clearly in educational measurements the variate may actually range from 0 to the full marks in the subject and generally the values are integral. But the assumptions of continuity and of unlimited range from $-\infty$ to $+\infty$ will not be meaningless. (By assumption of continuity we mean that a variate value x— an integer— is really the middle value of the class-interval $x-0.5$ to $x+0.5$). Let the distribution functions

* Presented at the 41st session of the Indian Science Congress, Hyderabad, 1954.

o x and y be $dF(x) = f(x)dx$, and $d\Phi(y) = \phi(y)dy$ respectively. (We are assuming that the distributions belong to the continuous type, which is supported by experience.) The percentile ranks p and p' say of two scores x_i and y_i in the two subjects are then defined by

$$\frac{p}{100} = \int_{-\infty}^{x_i} f(x)dx = F(x_i)\,;\ \frac{p'}{100} = \int_{-\infty}^{y_i} \phi(y)dy = \Phi(y_i). \qquad \ldots \quad (1)$$

two scores x_i and y_i will be said to be equivalent if

$$p = p', \text{ i.e. } F(x_i) = \int_{-\infty}^{x_i} f(x)dx = \Phi(y_i) = \int_{-\infty}^{y_i} \phi(y)dy.$$

Diagrammatically the method is equivalent to drawing ogives (or cumulative percentage graphs) for the two subjects on the same graph and taking two scores in the two subjects as equivalent when a straight line parallel to the horizontal axis cuts the two ogives at points whose abscissa are these two scores.

It easily follows that the different quantiles (percentiles etc.) of the x–distribution are equivalent to the corresponding quantiles of the y–distribution.

When the two distributions are normal this definition will lead to a very simple relation. Suppose that x and y are normal with means m_x and m_y, and s.d.'s σ_x and σ_y respectively . Then if two scores x_i and y_i are equivalent we must have

$$\frac{x_i - m_x}{\sigma_x} = \frac{y_i - m_y}{\sigma_y}.$$

This relation shows that two scores in two different subjects following normal distribution are equivalent when the corresponding standardised scores are equal. In this situation the x and y scores may be mutually converted by the relation

$$\frac{x - m_x}{\sigma_x} = \frac{y - m_y}{\sigma_y}. \qquad \ldots \quad (3)$$

The equation (2) gives a functional relationship between the equivalent scores. Such functional relationships may be termed as "equivalence relationships" and the corresponding curve may be called the "equivalence curve". This is really the equipercentile curve, i.e., the two co-ordinates of every point on it have equal percentile ranks. If both the scores have normal distributions the functional relationship is linear and the equivalence curve may then be termed the equivalence line. Its equation is

$$y = m_y + \frac{\sigma_y}{\sigma_x}(x - m_x).$$

But in cases of non-normal variation of one or both of the distributions, such a simple functional relationship cannot be established. We may, however, approximate to it by obtaining a polynomial of appropriate degree as a satisfactory fit by taking one of x or y (whichever may be taken as the standard subject) as the independent variable and the other as dependent.

The equivalence curve or some approximation to it is necessary for suitable conversion of scores as it avoids laborious calculation of the equivalent scores every time from the theoretical distribution.

Suppose that the distribution of x scores and y scores belong to Pearsonian types III and II respectively, that is to say, the distribution of x is given by

$$dF = c\left(1+\frac{x}{A}\right)^{p} e^{-\gamma x}dx, \ (-A < x < \infty)$$

(origin at mean) and that of y is given by

$$dF = k\left(1-\frac{y^2}{a^2}\right)^{m} dy, \ (-a < x \leqslant a) \text{ (origin at mean).}$$

Two scores x_i and y_i in subjects X and Y respectively will be equivalent if

$$\int_{-A}^{x_i} C\left(1+\frac{x}{A}\right)^{p} e^{-\gamma x}dx = \int_{-a}^{y_i} k\left(1-\frac{y^2}{a^2}\right)^{m} dy. \qquad \dots \quad (4)$$

(We may put the lower limit equal to $-\infty$ in both the integrals).

Let the common value of the integrals in (2) be $\frac{i}{100}$. For different values of i, we may determine the quantiles (usually percentiles, or deciles) x_i and y_i for the two distributions, and thus get two sets of values of x and y, say $x_{10}, x_{20}, \dots, x_{80}, x_{90}$ and $y_{10}, y_{20}, \dots, y_{80}, y_{90}$ where $x_i \equiv y_i$. This gives a table of equivalent scores.

The task of approximating the exact equivalence relationship by means of some polynomial of adequate degree is lightened if in the preparation of the above preliminary table of equivalent scores the scores of the standard subject are taken as equispaced, in order to facilitate the use of orthogonal polynomials.

3. Procedures with sample data

Our discussion above is based on the population distribution of scores, which is not known in practice. With sample scores therefore the problem of specification has to be solved by the customary methods of graduation. We may graduate the frequency distribution of scores in the sample against the most important of the known sets of curves—the Pearsonian system. Thereafter the procedure will be exactly similar to that described above.

If we like to deal only with the sample and not to go through the process of graduation and calculation of percentiles by using tables, two courses of procedure are there: (i) drawing all the sample ogives on the same graph paper, and reading off from it the scores in the standard subject equivalent to any score in any other subject, and (ii) calculating several percentile scores for each sample distribution, plotting for each pair of subjects (one being the standard) points with corresponding percentile scores as co-ordinates in a graph, and connecting them by a free-hand curve—the equipercentile curve—which will be used to convert scores in various subjects in terms of the standard.

When the equivalence curve is set up, the raw scores of different individuals in all subjects may be converted into the scores in the standard subject. Since now all the scores are measured in the same scale we can sum them to get a valid measure by means of which the individuals may be ranked or scaled properly.

4. An example

Scores of a random sample of 500 students have been obtained from the records of a scholastic test. Each candidate has got six scores, i.e. scores in six different subjects: Vernacular, English, Classical Language, History, Geography and Mathematics, the full marks and pass marks in which were different. For ease of comparison (and not as a theoretical necessity) all the scores were transformed so as to correspond to full marks 100 in each subject. Though the subjects Vernacular and Classical Language allowed of many alternatives, no attempt was made to distinguish between different combinations possible under these broad headings. Here exists thus an element of heterogeneity.

The object of the investigation is to construct equivalence relationships in the form of equipercentile curves for the purpose of scaling.

The frequency distributions of scores in different subjects are shown in Table 1. These frequency distributions revealed one remarkable feature. In scholastic tests where pass, fail or class is important, there are very irregular frequencies, unusually high or low in borderline classes. The reason is obvious. Because there is a pass mark or a class mark in each subject, the examiner is not free from bias and as a result of this there are more cases in particular intervals (in different subjects) than in others. This is one of the factors contributing to non-normality of the distributions of scores.

It was found that the appropriate Pearsonian types for fitting the distributions of scores were : normal for Vernacular, type II for English and Geography, and type III for classical language, history and mathematics. The moment-coefficients etc., for the different distributions are shown in the Tables 2 and 3.

The percentile ranks of scores 0, 5, 10, ..., 95, 100 in Vernacular were found out and are shown in col. (2) of Table 4. The corresponding scores in other subjects i.e. those with the same percentile ranks were determined. For this purpose, the tables of Incomplete Gamma function and Incomplete Beta function were used when the distribution in the other subject was of type III and type II respectively. Thus a table of equivalent scores giving the scores in the other subjects which are equivalent to marks 0, 5, ..., 95, 100 in Vernacular is obtained. These are shown in cols. (3)—(7) of Table 4.

The approximate relationships between scores in Vernacular and the equivalent scores in the 5 other subjects were determined (separately for the 5 subjects) by fitting orthogonal polynomials. Denoting scores in Vernacular, Mathematics, Classical Language, History, English and Geography by x, y_m, y_c, y_h, y_e, y_g respectively, the equivalence relationships were obtained as :

$$\left.\begin{array}{lll}
\text{Mathematics and Vernacular :} & y_m & = -17.8555 + 0.8714x + 0.0083x^2 \\
\text{Classical language and Vernacular :} & y_c & = 1.0514 + 0.6545x + 0.0085x^2 \\
\text{History and Vernacular :} & y_h & = -11.1393 + 0.6518x + 0.0088x^2 \\
\text{English and Vernacular :} & y_e & = -1.40088 + 0.25421x + 0.01622x^2 - 0.000097x^3 \\
\text{Geography and Vernacular :} & y_g & = 11.42080 + 0.46064x + 0.01135x^2 - 0.00008x^3
\end{array}\right\} \quad \ldots \quad (5)$$

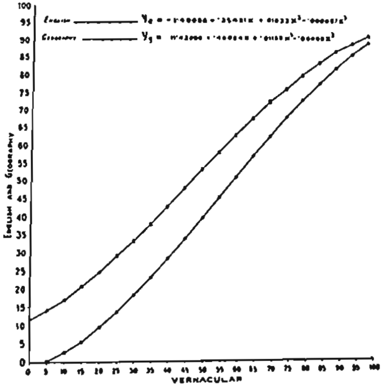
VERNACULAR
ENGLISH AND GEOGRAPHY

Fig. 1·

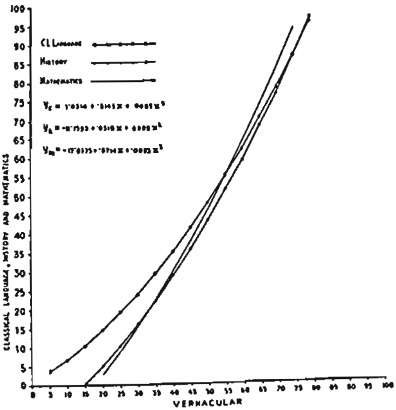
VERNACULAR
CLASSICAL LANGUAGE, HISTORY AND MATHEMATICS
CL. LANGUAGE
HISTORY
MATHEMATICS

Fig. 2

The above relationships are shown diagrammatically in Fig. 1 and Fig. 2. The fits provided were excellent except in the cases of English and Geography. This was due to some inaccuracies in the equivalent scores for these subjects (corresponding to scores 45 and 50 in Vernacular) and could be attributed to linear interpolation in portions of Incomplete Beta Function table. Entering the table for $q=5$ and $p=9.5$, 10 and 10.5 it was noticed that for $x=0.99$ the value of $I_x(p,q)$ are 0.6602641, 0.6579282 and 0.6498437 respectively and for $x=1$, $I_x(p,q)=1$. The table does not provide the value of I at some finer interval for the x-value in the range $0.99<x<1$, which is necessary in order to read the value of x for any $I_x(p,q)$ falling in the long range from nearly 0.65 to 1. For these two subjects the third degree polynomials did not give very good results, while for the others 2nd degree curves were very good.

On the basis of these relationships another set of equivalent scores was again prepared [cols. (8)-(12), Table 4]. Either the relationships themselves or the 5 equivalence curves drawn may be used for finding scores equivalent to given scores in Vernacular.

The method was applied to examine the ranking by the present system (i.e. using the total of raw scores) of the 15 top-ranking candidates i.e. those securing highest total ordinary scores in the sample. All scores were converted to equivalent scores in Vernacular; these converted scores were added and used for scaling the same candidates (Table 5). The two scales are given as follows :

Old ranking :—1, 2, 3, 4, 4, 6, 7, 8, 8, 10, 10, 10, 13, 14, 14.

New ranking :—3, 2, 1, 4, 5, 7, 6, 14, 7, 10, 11, 9, 12, 12, 15.

A comparison of the old ranking and this new ranking is illuminating. It is a pointer to the fact that we err seriously by the old method.

REFERENCES

GREENALL, P. D. (1949): The concept of equivalent scores in similar tests. *British Journal of Psychology*, Statistical Section, 2, 30-40.

HUSSAIN, Q. M. (1941): Standardisation of examination marks. *Sankhyā*, 5, 295-300.

MAHALANOBIS, P. C. and CHAKRAVARTI, K. N. (1934): Studies in educational tests No. 3., *Sankhyā*, 1. 245-266.

TABLE I. FREQUENCY DISTRIBUTIONS OF SCORES IN SIX SUBJECTS OBTAINED BY 500 STUDENTS

score	frequency					
	Vernacular	English	Classical Language	History	Geography	Mathematics
0— 4		3		4	1	9
5— 9		6	2	9	0	8
10—14		12	7	12	4	18
15—19	6	23	9	29	7	13
20—24	7	35	12	28	12	32
25—29	18	45	8	15	2	4
30—34	34	74	80	118	69	101
35—39	56	72	79	58	45	58
40—44	84	78	73	52	53	51
45—49	74	53	62	40	60	46
50—54	104	46	55	42	61	42
55—59	53	29	33	26	36	31
60—64	36	18	21	25	48	21
65—69	16	5	19	13	32	14
70—74	9	1	20	9	44	19
75—79	0		8	5	18	16
80—84	3		9	14	6	6
85—89			2	1	2	5
90—94			1			4
95—99						2
total	500	500	500	500	500	500

TABLE 2. MOMENTS, β_1 AND β_2 FOR THE DISTRIBUTIONS OF SCORES IN 6 SUBJECTS

subjects	Vernacular	Mathematics	Classical Language	History	English	Geography
mean	47.07	42.57	45.14	40.14	38.87	49.94
variance *	5.121804	14.045004	8.761616	11.221616	6.863724	9.722256
μ_3 *	− 0.405118	17.989128	11.677442	15.100802	− 2.394132	− 1.597358
μ_4 *	82.658661	600.247617	250.894092	387.000413	128.824563	259.552977
β_1	0.001222	0.116803	0.202740	0.103304	0.017726	0.002776
β_2	3.150959	3.042890	3.268209	3.074060	2.734510	2.745940

* (Class-interval)^2, (Class-interval)^3 and (Class-interval)^4 being the units of measurements for Variances, μ_3 and μ_4 respectively.

TABLE 3. PEARSONIAN CURVES FITTING THE SIX FREQUENCY DISTRIBUTIONS

subject	type	equation	values of parameters			
Vernacular	normal	$\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{x^2}{2\sigma^2}}$ $-\infty < x < \infty$	mean = 47.07			σ = 11.315920
English	II	$y_0\left(1-\frac{x^2}{a^2}\right)^m$ $-a < x < +a$	mean = 38.87		m = 8.709861	a^2 = 3534.765900
Geography	II		mean = 49.94		m = 9.308234	a^2 = 2627.010550
Classical Language	III	$y_0\left(1+\frac{x}{A}\right)^p e^{-\gamma x}$ $-A < x < \infty$	mean = 45.14	γ = 0.300121	p = 18.729703	A = 65.739162
History	III		mean = 40.14	γ = 0.295484	p = 23.494195	A = 82.895165
Mathematics	III		mean = 42.57	γ = 0.312206	p = 33.188034,	A = 109.473173

origin at mean

TABLE 4. EQUIVALENT SCORES BASED ON EQUATION 2 AND ON REGRESSION EQUATIONS 5

		equivalent scores (equation 2)					equivalent scores (equations 5)				
Vernacular	percentile rank	Classical Language	History	Mathematics	English	Geography	Classical Language	History	Mathematics	English	Geography
0	0.00150					10.88	1				11
5	0.01	3.55				13.82	4			0	14
10	0.05	6.90				16.88	7			3	17
15	0.23	10.83	0.46		3.04	20.54	11	1		6	21
20	0.84	15.00	5.40	2.83	7.63	24.44	15	5	3	9	25
25	2.56	19.49	10.73	9.11	12.60	28.67	19	11	9	14	29
30	6.57	24.24	16.41	15.79	17.94	33.17	24	16	16	18	33
35	14.31	29.63	22.54	22.86	23.56	37.90	29	22	23	23	38
40	26.61	34.64	29.07	30.30	29.40	42.81	35	29	30	29	43
45	42.74	41.39	36.01	38.15	33.98	46.62	41	36	38	34	48
50	60.21	47.90	43.40	46.42	42.49	53.87	48	43	46	40	53
55	73.83	54.86	51.23	55.11	47.41	57.06	55	51	55	46	58
60	87.34	62.26	59.51	64.21	53.17	62.86	63	60	64	51	63
65	94.35	70.14	68.27	73.78	58.75	67.51	70	68	74	57	67
70	97.86	78.44	77.47	83.77	64.01	71.96	79	78	84	63	72
75	99.32	87.26	87.19	94.18	68.93	76.14	87	87	94	68	76
80	99.82	96.50	97.41		73.42	79.97	97	97		73	80
85	99.96				77.52	83.47				78	83
90	99.99				80.53	86.06				82	86
95	99.99886				84.25	89.44				86	89
100	99.99985				86.83	91.70				89	91

TABLE 5. ORIGINAL AND EQUIVALENT SCORES AND RANK OF STUDENTS SECURING HIGHEST TOTAL ORDINARY SCORES FROM THE SAMPLE OF 500.

students	Vernacular	Classical Language	History	Mathematics	English	Geography	total	rank
	actual score							
1	71	90	71	85	71	64	452	1
2	71	70	78	92	62	78	451	2
3	73	62	71	75	67	86	434	3
4	64	68	80	77	67	72	428	4
5	64	74	80	83	59	68	428	4
6	84	81	84	73	59	38	419	6
7	63	65	54	97	56	80	415	7
8	58	60	83	88	56	62	407	8
9	66	71	55	84	57	74	407	8
10	62	81	55	72	62	70	402	10
11	60	80	48	80	62	72	402	10
12	62	58	58	82	68	74	402	10
13	64	76	60	61	62	70	393	13
14	67	58	64	58	67	74	388	14
15	61	71	44	80	62	70	388	14
	equivalent scores (Vernacular equivalent)							
1	71	77	67	71	78	61	425	3
2	71	65	70	74	70	78	428	2
3	73	60	67	66	74	90	430	1
4	64	64	72	67	74	70	411	4
5	64	68	72	70	67	66	407	5
6	84	72	74	65	67	35	397	7
7	63	62	57	76	64	80	402	6
8	58	58	73	73	64	59	384	14
9	66	65	58	70	65	73	397	7
10	62	72	58	64	70	68	394	10
11	60	71	58	67	70	70	391	11
12	62	57	59	69	75	73	395	9
13	64	69	60	59	70	68	390	12
14	67	57	63	56	74	73	390	12
15	61	65	50	67	70	68	381	15

Paper received : September, 1953.

ON THE EVALUATION OF GROSS VALUE IN AGRICULTURE BY MAKING USE OF DISTRICTWISE PRICE DATA AND THE STATE OUTTURN

By M. MUKHERJEE

Central Statistical Organization, New Delhi

1. Usually it happens that at the time of estimation of the gross value in agriculture, districtwise price data are available, but the outturn is available for the State as a whole. The purpose of this paper is to find out theoretically an appropriate price average for the evaluation of the State outturn in this case and then to test the theoretical findings with the districtwise data on both production and prices available in respect of past period for a number of States. The work is important because agriculture contributes roughly half of the national income in India, and even a small addition to accuracy in the estimation of gross value in this sector would materially reduce the overall error of our national income estimate.

2. The problem may be posed as follows. Consider a State with n districts. The prices (p_i) in n districts are known. The production figures (w_i) are not known, but $\sum_{i=1}^{n} w_i = W$ is known. $\Sigma w_i p_i$ gives theoretically the most satisfactory estimate of the gross value. The problem is to estimate $\xi = f(p_i)$, such that $W\xi$ is in some sense the best possible estimate for $\Sigma w_i p_i$.

3. Let
$$\sigma_w^2 = \frac{\sum_{i=1}^{n} (w_i - \bar{w})^2}{n}, \quad \bar{w} = \frac{\sum_{i=1}^{n} w_i}{n} = \frac{W}{n},$$

$$\sigma_p^2 = \frac{\sum_{i=1}^{n} (p_i - \bar{p})^2}{n}, \quad \bar{p} = \frac{\sum_{i=1}^{n} p_i}{n}.$$

$$r_{wp} = \frac{\sum_{i=1}^{n} (w_i - \bar{w})(p_i - \bar{p})}{n\sigma_w \sigma_p} = \frac{\sum_{i=1}^{n} w_i p_i - n\bar{w}\bar{p}}{n\sigma_w \sigma_p} = \frac{\Sigma w_i p_i - W\bar{p}}{n\sigma_w \sigma_p}.$$

4. If follows from paragraph 3, that if on *a priori* grounds r_{wp} can be assumed to be zero, then, $\Sigma w_i p_i$ is estimated best when $\xi = \bar{p}$. When ξ equals the median, and there is no reason to believe that the distribution of p's is skew, the median will be an equally good estimator. If, however, p's have a distribution with positive skewness, the median will be under-estimating; but if the skewness is negative, the median will be over-estimating. In either case, mean will be a better estimator than the median. In 21 cases studied by us, the mean is greater than the median in 15 cases (see Table 1). Thus it is likely that positive skewness exists in a larger number of cases. If this is so, then on the assumption of zero correlation, mean becomes a better estimator than median.

5. On *a priori* grounds, some negative correlation may, however, be expected between production and prices. Where there is surplus production, prices are likely to be low. It may be reasonably supposed that districts with larger production have on an average a greater chance of being surplus than districts with smaller production. Quite possibly some

district with small production will be surplus, but still the above supposition may generally hold. Under this supposition there will be a negative correlation between production and prices. In our study, in 15 cases out of 21, we have found negative correlation, lending support to the above contention (see Table 1).

6. Now if $r_{wp} < 0$, and no hypothesis is made as regards skewness of p's, it follows from paragraph 3 that $\bar{p}$ will be over-estimating. In this case the geometric (or harmonic) mean may be a better estimator than $\bar{p}$.

7. The available evidence however leads to the hypothesis that both $r_{wp} < 0$ and skewness is positive. In this case since the median $< \bar{p}$, it is likely to be a better estimator than the mean.

8. Let us call $\Sigma w_i p_i / W = \mu$ and geometric (or harmonic) mean m_g. In the case discussed in paragraphs 6 and 7, the geometric mean (or harmonic mean) will be better estimators only if the excess of $\bar{p}$ over μ is greater than the excess of μ over m_g, when μ is greater than m_g.

9. A detailed study of the above problem has been made in respect of the median m. By our hypothesis, $\bar{p} > m$, and $\bar{p} > \mu$ since $r_{wp} < 0$. Now if $m > \mu$, it will be a better estimator than the mean. If $m < \mu$, but $\mu - m < \bar{p} - \mu$, then also median will be better. When $\mu - m = \bar{p} - \mu$, $\bar{p}$ and m will be equally good. But when $\mu - m > \bar{p} - \mu$, mean will be a better estimator than the median.

10. However, since

$$\frac{\mu - m}{\mu - \bar{p}} = 1 + \frac{\bar{w}}{\sigma_w} \cdot \frac{\bar{p} - m}{\sigma_p} \cdot \frac{1}{r_{wp}},$$

the magnitude of under-estimation by the use of median will be less than the magnitude of over-estimation by the use of mean so long as $\frac{\bar{w}}{\sigma_w} \cdot \frac{\bar{p} - m}{\sigma_p} \cdot \frac{1}{\gamma_{wp}}$ is numerically less than 2. As districtwise production figures are expected to show a large coefficient of variation, and $\frac{\bar{p} - m}{\sigma_p}$ is necessarily a very small quantity, the above expression can exceed 2 only when r_{wp} is very small, but this case is precluded by the hypothesis of real negative correlation. Hence, the case outlined at the end of paragraph 9 has a very small chance of occurrence for the type of data we are dealing with and we may conclude that when $r_{wp} < 0$ and skewness is positive (i.e., in the normal case in respect of the data we are using), median is a better estimator than the mean. This is corroborated by a larger study of 83 cases for which median gives an overall under-estimation of 0.6 p.c. while mean gives an overall over-estimation of 1.5 p.c.

11. If we go back to the position of paragraph 6, and remove the assumption of positive skewness, since considerations in paragraph 10 show that the magnitude of over-estimation by the use of mean is likely to be greater than the magnitude of under-estimation by the use of median, when μ lies in the range $(\bar{p}, m)$, it follows that when $\bar{p} > m$, m is a better estimator than $\bar{p}$. If, however, $m > \bar{p}$, μ being obviously less than $\bar{p}$, $\bar{p}$ is a better estimator than m in this case. Thus, we arrive at the interesting rule that when $r_{wp} < 0$ and no assumption is made regarding skewness, the smaller of the mean and median is the better estimator. The rule is useful because this can be followed in practice.

12. For positive correlation, on the other hand, it can be easily shown that the larger of the mean and median is the better estimator.

13. The experimental results are presented in Table 1 below. It will be seen that the rules given in paragraphs 11 and 12 have been violated only once (in the case of Hyderabad, Jowar). In this case values of $\frac{\bar{w}}{\sigma_w}$ and $\frac{\bar{p}-m}{\sigma_p}$ are both found to be unexpectedly large, and the result is more due to this than due to the smallness of the correlation coefficient.

TABLE 1. PRICE AVERAGES AND CORRELATIONS BETWEEN PRICE AND PRODUCTION

crop	state	year	no. of districts	weighted mean	mean	median	r_{wp}
winter rice	Pihar	1947–48	15	18.3000	18.2300	16.0000	+0.0557
autumn rice	"	"	15	14.5000	16.7000	15.0000	−0.2520
wheat	"	"	13	22.0400	21.9000	22.0000	+0.0207
barley	"	"	11	12.1000	12.9000	12.8000	−0.3400
gram	"	"	15	14.0083	15.1000	13.1000	−0.1062
gur	"	"	15	14.6000	16.6000	16.0000	−0.3864
Indian corn	"	"	12	10.2000	10.4000	10.0000	−0.1130
rice	U.P.	1945–46	31	17.7000	18.4000	17.8000	−0.2170
"	"	1946–47	38	17.7800	17.6600	17.7500	+0.0380
"	"	1947–48	43	19.8000	19.2500	17.7500	+0.2170
"	"	1948–49	43	26.2000	26.6000	26.7000	−0.1230
wheat	East Punjab	1948–49	12	14.4000	14.6200	14.3000	−0.5000
gram	" "	"	12	8.5000	8.0300	8.5000	−0.2960
groundnut	Hyderabad	"	14	62.7000	61.3000	62.0000	+0.3690
jowar	"	"	12	39.6000	40.5000	33.5000	−0.1340
rice	Bombay	"	18	0.6800	0.7700	0.7400	−0.2988
wheat	"	"	14	0.6000	0.5640	0.4750	+0.3739
gram	"	"	18	0.4800	0.5590	0.5350	−0.5059
gur	"	"	18	0.6400	0.6530	0.6800	−0.1112
bajra	"	"	14	0.3900	0.4240	0.4000	−0.0013
jowar	"	"	16	0.3800	0.3056	0.3650	−0.1610

Paper received: March, 1951.