EAGLE
REGD

Office File

SUPERIOR

Previous File No.

ON A NEW SIRE INDEX FOR MILK PRODUCTION

K. S. Krishnan

Dissertation submitted in partial fulfilment of the requirements for the award of diploma in Agricultural and Animal Husbandry Statistics of the Indian Council of Agricultural Research,

New Delhi

1957

043:60

ON A NEW SIRE INDEX FOR MILK PRODUCTION

(K. S. Krishnan, ICAR)

1. Introduction

In any scientific programme for improvement of dairy cattle greatest attention should be paid to the selection of breeding bulls in view of the large number of progenies that can be raised by a bull in his life time as compared to a cow. With the adoption of the artificial insemination technique the problem has received added importance as the number of progenies that can be obtained from a bull by following this technique is increased manifold.

Milk yield being a sex limited character the phenotypic expression of which is confined to female sex only, selection of bulls on its performance is ruled out. One alternative possible is to take as the criterion for selection a character which is highly correlated with milk yield but not sex limited. Available evidence seems to indicate little possibility of such a character being found. Selection based on body conformation is an attempt in this direction and this has rarely resulted in bringing about the desired improvement.

An alternative and more promising method is to base the selection of bulls on the performance of his female relatives. The performance of the daughters of a bull will be more useful in judging its breeding worth than that of other female relatives since the information provided by the latter are limited (Lush 1949).

An estimate of the breeding worth of a sire is termed sire index. Different sire indices have been proposed in the past. In the course of extensive analysis of breeding data undertaken in the Council it has been observed that the indices in common use are subject to certain limitations. In the present thesis these limitations have been examined and a sire index relatively free from them has been proposed. The

-/-

relative merits of the new index vis a vis the common ones has been empirically demonstrated on the data pertaining to helf a dozen different herds.

2. Review of literature

Ways of estimating a sire's breeding worth from the performance of his progeny have been studied extensively for dairy characters. Perhaps the earliest attempt to formulate an index expressing the breeding value of a sire was ~~perhaps~~ that of Hansson (1913). The 'Intermediate' or 'Equal-Parent' index was suggested by Yapp (1925) which rests on the assumption that the genetic value of the daughter is mid-way between the average production of dam's and sire's potential transmitting ability. Wright (1931) suggested a modification over the intermediate index, combining information about a bull's dam and (m-1) of his full sisters into a single index.

In order to compare the sires used in different herds Von Patow (1930) and Krüger (1938) have advocated the use of ~~expressing~~ records expressed as deviations from the contemporary herd average, in order to correct for general environmental conditions applying for some herds but not to all. This in effect assumes that all differences between herd averages are environmental. This assumption is avoided in the 'Regression index' proposed by Rice (1944) for the inter-herd comparisons. The Regression index simply regresses the intermediate index halfway back to the breed average which is the simple average of all the cows of the particular breed under consideration maintained in different herds. Berge (1944) suggested a modified index on the assumption that the regression coefficient of daughter on dam is of the order of 3/8. A method of correcting the individual lactation records for the inequality in lactation period was suggested and the procedure for calculating the intermediate index and its standard error from these records indicated by Sukhatme (1944).

3. Indices in common use and their shortcomings.

Nearly all the indices proposed for sire evaluation are based on simple daughter average or daughter-dam differences or a combination of the two. The simple daughter average index which lays the emphasis entirely on the former and the intermediate index which is simply the daughter average plus the amount by which the daughters exceed their dams are the two common indices widely in use for sire testing in breeding herds.

The simple daughter average index D of a sire is defined as

$$D = \bar{y} \qquad (1)$$

where $\bar{y}$ is the simple average of its daughters with records.

The intermediate index I of a sire is defined as

$$I = 2\bar{y} - \bar{x} \qquad (2)$$

where $\bar{y}$ and $\bar{x}$ are respectively the average of all its daughters with records and the average of the dams of these daughters.

Any method of sire evaluation is likely to be affected by the culling of female calves and heifers. No satisfactory procedure is available for overcoming this defect in the estimation of the sire index. Retention of all female progeny till the completion of their first lactation, except for reasons of disease or confirmed infertility, is therefore, essential for sire evaluation.

Apart from this common limitation, the simple daughter average index does not take into account possible assortive matings resulting in unequal average production levels of the dams sired by different bulls. One limitation of the Intermediate index is that it overcorrects for the differential level of production of dams mated to different sires; that is if the set of cows mated to a sire is inferior to the average, the index over-estimates the sire's breeding worth whereas if the dams are above a ... - ... that is worth. Secondly, the inter- ... i ... error, it the consequence that ... discriminate between even the best and the ... bulls that may be used on a herd.

...

... corrected index is obtained from the formula:

$$S = \bar{y} - b(\bar{x} - A) \qquad (3)$$

... $\bar{y}$ is the average for the daughters of the sire, $\bar{x}$ is the ... the dams of these daughters, b is the intrasire regression of the daughters' performance on that of its dam, and A ... average.

This index is the regression estimate familiar in sampling theory, ... performance is treated as the auxilary variable. It may be termed the corrected daughter average index or ... for short.

The corrected index, as will be shown later, is free from the bias due to the unequal levels of production of the cows mated to different bulls. Further the corrected index is at least four times as efficient as the intermediate index for the ... ability of ... percent or less. Even when the value of the heritability ... than fifty percent the corrected index is seen to be superior ... to a lesser degree. The corrected index which is unbiassed ... efficiency in no case lower than that of the simple daughter average index.

~~The theory relating to the genetic basis of the three indices, viz the simple daughter average index~~.

The theory relating to the genetic basis of the three indices, viz the simple daughter average index the intermediate index and the corrected index is given and their expectation, sampling variances and relative efficiencies are worked out in the following sections.

5·1 Genetic model

Let x_1, x_2, ..., x_N be the phenotypic value (... basis, such as the first lactation yield say) of ... belonging to a herd under the set of conditions of man... ...vir... obtaining in a giv n farm.

Let $A = \frac{x_1 + x_2 + x_N}{N}$ be the av r ge level of production of the...

Let $g_1, g_2, \ldots, g_N$ be the deviations in the production level of ... from A which can be ascribed to genotypes of these ... ...t $\delta_1, \delta_2, \ldots, \delta_N$ the corresponding environmental ...s.

Then ... y $x_l = A + g_l + \delta_l$ (4)

under ...ption that interaction between genotype and environment is ...nt.

Let ...b et n_i of these cows are mated to the i-th sire (i = 1,2,......k). The sub set is not necessarily chosen at random. It will be assumed that each cow has only one d ughter with records. This, however, may not be r alised in actual practice, the consequence of which will be discusse in a later section.

... $x_{i1}, x_{i2}, \ldots, x_{in_i}$ be the phenotypic values ... these cows, $g_{i1}, g_{i2}, \ldots, g_{in_i}$ the genotypic ...ations and $\delta_{i1}, \delta_{i2}, \ldots, \delta_{in_i}$ the corresponding environmental deviations.

Then $x_{ij} = A + g_{ij} + \delta_{ij}$ (5)

Let $\bar{g}_i = \frac{\sum_j g_{ij}}{n_i}$ and $\bar{\delta}_i = \frac{\sum_j \delta_{ij}}{n_i}$

$E(\bar{g}_i)$ ... ($\bar{\delta}_i$) where E denotes the expectation under a giv... ...st ... may not vanish unl... the cows are allotted to di... ...sires at random.

Let g'_i be the genotypic deviation of the i-th sire from A. Then under the assumption that dominance and generic interaction are ...t, ... genotypic deviation of the j-th daughter of the i-th sire ... $\frac{g'_i + g_{ij}}{2}$

Let y_{ij} be the phenotypic value of the yield of this daughter and e_{ij} ... environmental component of this yi l...

Then $y_{ij} = A + \frac{g_i' + g_{ij}}{2} + e_{ij}$ (6)

If the differences in the milk yield of two daughters which is attributable to non-heritable causes are purely of a random nature, then e_{ij} and $e_{ij'}$, the environmental deviations in the yields of two daughters of the same bull, are uncorrelated and the expected value of e_{ij} will be the same for the daughters of every sire. e_{ij} may be so defined that this expectation $E(e_{ij})$ is zero.

With the set up so defined the expected values of the three indices may next be examined.

5.2. Simple daughter average index.

The simple daughter average index D_i of a sire can easily be seen to be

$$\begin{aligned} D_i = \bar{y}_i &= A + \frac{g_i' + \bar{g}_i}{2} + \bar{e}_i \\ &= A + \frac{g_i'}{2} + \frac{\bar{g}_i}{2} + \bar{e}_i \end{aligned} \quad (7)$$

$E(\bar{e}_i) = 0$ but $E(\bar{g}_i) \neq 0$ unless every cow has an equal chance of being mated to any bull. If this condition is satisfied, then the ~~bias~~ deviation introduced by $\frac{\bar{g}_i}{2}$ in D_i, as an estimate of the breeding worth of the sire, is of the nature of a sampling error similar to $\bar{e}_i$ and only lowers the precision of the estimate.

In practice these conditions may not be fulfilled on account of (i) possible phenotypic assortive mating based on milk yield or related characters and (ii) inbreeding which is nothing but genotypic assortive mating. Hence in the further development of the theory $E(\bar{g}_i)$ has not been assumed to be necessarily zero but to have a value γ_i which may or may not be equal to zero. ~~The daughter average index is also correlated with the average of their dams with $\bar{x}_i$ as is shown below where~~

The daughter average index D_i is also correlated with the average of their dams $\bar{x}_i$ as is shown below where $\bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}$.

$$E(\bar{x}_i) = A + \gamma_i \quad (8)$$

Then

$$E(D_i) = E(\bar{y}_i)$$

$$= E\left\{A + \frac{g_i' + \bar{g}_i}{2} + \bar{e}_i\right\}$$

$$= A + \frac{E(g_i')}{2} + \frac{\gamma_i}{2} \quad (8a)$$

$$\mathrm{Cov}(D_i, \bar{x}_i) = \mathrm{Cov}(\bar{y}_i, \bar{x}_i)$$

$$= E\left[\left\{\bar{x}_i - E(\bar{x}_i)\right\}\left\{\bar{y}_i - E(\bar{y}_i)\right\}\right]$$

$$= E\left[\left\{(A + \bar{g}_i + \bar{\delta}_i) - (A + \gamma_i)\right\}\left\{\left(A + \frac{g_i' + \bar{g}_i}{2} + \bar{e}_i\right) - \left(A + \frac{E(g_i')}{2} + \frac{\gamma_i}{2}\right)\right\}\right]$$

$$= E\left[\left\{(\bar{g}_i - \gamma_i) + \bar{\delta}_i\right\}\left\{\tfrac{1}{2}(\bar{g}_i - \gamma_i) + \tfrac{1}{2}\left(g_i' - E(g_i')\right) + \bar{e}_i\right\}\right]$$

$$= \tfrac{1}{2} V(\bar{g}_i) + \tfrac{1}{2}\mathrm{Cov}(\bar{g}_i, g_i')$$

$$+ \tfrac{1}{2}\mathrm{Cov}(g_i', \bar{\delta}_i) + \tfrac{1}{2}\mathrm{Cov}(\bar{g}_i, \bar{\delta}_i)$$

$$+ \mathrm{Cov}(\bar{g}_i, \bar{e}_i) + \mathrm{Cov}(\bar{\delta}_i, \bar{e}_i) \quad (9)$$

Now

$$\tfrac{1}{2} V(\bar{g}_i) = \frac{1}{2}\frac{\sigma_g^2}{n_i}$$

where σ_g^2 is the genetic variance in dams. The term in cov $(\bar{g}_i, g_i')$ is non-zero if there is association between the average genotype of the cows mated to a sire and the genotype of the sire itself. This term will be positive if there is inbreeding or mating of like to like and negative if mating of unlike individuals is preferred. Under a system of random mating or out-crossing the term will vanish.

Cov (g_i' , $\bar{\delta}_i$) is the covariance between the genotype of a sire and the deviation in the average production of the dams due to environment. When, however, no preferential environment is given to the mates of any particular bull, this term will vanish.

The term cov ($\bar{g}_i$, $\bar{\delta}_i$) $\neq 0$ if (i) groups of cows of different production capacities are assigned to different bulls and (ii) these groups are given different treatments as regards feeding and management. Similarly the term cov ($\bar{g}_i$, $\bar{e}_i$) will be non-zero if (i) cows of different production capacities are mated to different bulls and (ii) the daughters of the different groups are reared differently. Both the above terms will vanish under random allotment of cows to different bulls. Even if this is not so the first will vanish if the conditions of management of the mates of different bulls is not related to their production capacities and the second if all the daughters born are reared under uniform conditions.

The last term in (9), viz cov($\bar{\delta}_i$, $\bar{e}_i$) is attributable to environmental correlation between the daughters of a bull and their dams and will not vanish if the daughters of better managed cows are reared better.

Each one of the variance and the covariances in the right hand side of equation (9) will contribute for the correlation of D_i and $\bar{x}_i$. If inbreeding or phenotypic assortive mating is not deliberately practised and environmental deviations for daughters and dams are uncorrelated, correlation between D_i and $\bar{x}_i$ can readily be seen to be

$$\frac{\frac{1}{2n_i}\sigma_g^2}{\sqrt{V(\bar{x}_i)\cdot V(\bar{y}_i)}}$$

$$= \frac{\frac{1}{2n_i}\sigma_g^2}{\frac{1}{n_i}\sqrt{\sigma_x^2\cdot\sigma_y^2}}$$

$$= \frac{\sigma_g^2}{\sqrt{\sigma_x^2\cdot\sigma_y^2}} \qquad (10)$$

where σ_x^2 and σ_y^2 are the variances between the mates of a bull and its daughters respectively averaged over all the bulls. Since $\frac{\sigma_g^2}{\sigma_x^2}$ is the heritability coefficient h^2 among the dams, correlation between D_1 and $\bar{x}_1$ may be written as $\frac{1}{2}h^2\sqrt{\sigma_x^2/\sigma_y^2}$

The daughter average index is thus seen to be positively correlated with the average level of production of the cows mated to the sire. This implies that the breeding worth of the sire is over-estimated if the cows allotted to a sire happen to be superior on the whole and is under-estimated if an ~~lot of~~ inferior lot of cows is ~~cows are~~ given to the bull.

When it is desired to compare two sires s_1 and s_2 in respect of their transmitting abilities, and if the simple daughter average index is used for this purpose, then

$$D_1 - D_2 = \bar{y}_1 - \bar{y}_2$$

$$= \frac{g_1' - g_2'}{2} + \frac{\bar{g}_1 - \bar{g}_2}{2} + \bar{e}_1 - \bar{e}_2 \qquad (11)$$

($\bar{e}_1 - \bar{e}_2$) is the difference attributable to environmental causes. This difference will be in the nature of an error if the deviations ascribable to non-heritable causes are of a random nature but a source of bias otherwise. Similarly if the chance of mating a cow to Sire 1 is the same as that of mating the same cow to Sire 2, then ($\bar{g}_1 - \bar{g}_2$) will be of the nature of an error and hence ($D_1 - D_2$) can be taken as a valid estimate of ($\frac{g_1' - g_2'}{2}$) which is half the difference between the genotypic values of the two bulls. When phenotypic assortive mating or inbreeding is practised ($\bar{g}_1 - \bar{g}_2$) may not vanish and hence ($D_1 - D_2$) is a biassed estimate of $\frac{g_1' - g_2'}{2}$

§3 X. Intermediate index

The intermediate index I_i for the sire i is

$$I_i = 2\bar{y}_i - \bar{x}_i$$

$$= 2\left\{A + \frac{g_i' + \bar{g}_i}{2} + \bar{e}_i\right\} - \left\{A + \bar{g}_i + \bar{\delta}_i\right\}$$

$$= A + g_i' + 2\bar{e}_i - \bar{\delta}_i \qquad (12)$$

$E(\bar{e}_i) = 0$ but $E(\bar{\delta}_i)$ may not vanish unless the conditions of feeding and management under which the records of dams are made can be taken to be randomly distributed in so far as the different groups of cows are concerned

$$Cov(I_i, \bar{x}_i) = Cov\left\{(A + g_i' + 2\bar{e}_i - \bar{\delta}_i), (\bar{g}_i, \bar{\delta}_i)\right\}$$

$$= -V(\bar{\delta}_i) + Cov(\bar{g}_i, g_i') + Cov(g_i', \bar{\delta}_i)$$

$$- Cov(\bar{g}_i, \bar{\delta}_i) + 2\,Cov(\bar{g}_i, \bar{e}_i) - 2\,Cov(\bar{\delta}_i, \bar{e}_i) \qquad (13)$$

The implications of the different covariance terms in (13) are the same as in (9) discussed in the previous section. Even if the assumptions needed for these covariances to vanish hold, the term $-V(\bar{\delta}_i)$ will remain and I_i and $\bar{x}_i$ will therefore be correlated.

The intermediate index is therefore negatively correlated with the average level of production of the cows mated to the sire with the consequence that the breeding worth of the sire is overestimated if the cows allotted to a sire are high yielders and is under-estimated if poor yielders are mated to the bull.

When the breeding values of two sires are ~~and~~ ~~are~~ compared by using the intermediate index, then

$$I_1 - I_2 = (A + g_1' + 2\bar{e}_1 + \bar{\delta}_1) - (A + g_2' + 2\bar{e}_2 - \bar{\delta}_2)$$

$$= (g_1' - g_2') + 2(\bar{e}_1 - \bar{e}_2) - (\bar{\delta}_1 - \bar{\delta}_2) \qquad (14)$$

The expectation of the term $(\bar{e}_1 - \bar{e}_2)$ will vanish and the term will contribute to error if the environmental conditions under which the records of daughters are made can be taken to be random. If however preferential treatment is given to the daughters of one bull as against those of the other, the term will introduce bias in the comparison of the two bulls. The situation with regard to the term $(\bar{\delta}_1 - \bar{\delta}_2)$, which refers to the environment of the dams, is precisely similar to that of $(\bar{e}_1 - \bar{e}_2)$.

5·4. Corrected index

The corrected index S_i for the sire i is given by

$$S_i = \bar{y}_i - b(\bar{x}_i - A) \qquad (15)$$

where b is the intrasire regression of daughter's yield on that of its dams.

The corrected index S_i makes an allowance for the bias present in the daughter average index due to the differential level of production of the dams mated to different sires. The appropriateness of the adjustment can be seen from the following argument.

It can easily be shown that the linear regression coefficient of genotypic value on the phenotypic value is equal to the heritability coefficient, h^2. Hence

$$g_{ij} = h^2(x_{ij} - A)$$

$$\therefore \bar{g}_i = h^2(\bar{x}_i - A).$$

As shown in (8a), $E(D_i)$ contains the term $\frac{\gamma_i}{2}$ and its effect will be in the nature of a bias unless every cow has an equal chance of being mated to a bull. Where this condition is not satisfied, the addition of the term $-\frac{1}{2}h^2(\bar{x}_i - A)$ made to D_i i.e. to $\bar{y}_i$ adjusts for the effect of the deviation due to $\frac{\gamma_i}{2}$ and renders the resulting estimate unbiassed. It is known that twice the intrasire regression of daughter on dam is the most efficient estimate of heritability (h^2). Hence S_i will be an estimate of $A + \frac{g_i'}{2}$ free from the bias due to unequal production levels of dams.

Unlike D_i and I_i, the index S_i is not correlated with the dams' average. This can be seen as follows:

$$Cov(S_i, \bar{x}_i) = Cov\left[\left\{\bar{y}_i - b(\bar{x}_i - A)\right\}, \bar{x}_i\right]$$

$$= Cov(\bar{y}_i, \bar{x}_i) - Cov(b\bar{x}_i, \bar{x}_i) + Cov(Ab, \bar{x}_i) \quad (16)$$

From equation (9)

$$Cov(\bar{y}_i, \bar{x}_i) = \frac{1}{2n_i}\sigma_g^2 + \frac{1}{2}Cov(\bar{g}_i, g_i')$$

$$+ \frac{1}{2}Cov(g_i', \bar{\delta}_i) + \frac{1}{2}Cov(\bar{g}_i, \bar{\delta}_i)$$

$$+ Cov(\bar{g}_i, \bar{e}_i) + Cov(\bar{\delta}_i, \bar{e}_i)$$

$$Cov(b\bar{x}_i, \bar{x}_i) = E(b\bar{x}_i^2) - E(b\bar{x}_i).E(\bar{x}_i)$$

$$= E(b)\,E(\bar{x}_i^2) - E(b)\left\{E(\bar{x}_i)\right\}^2$$

$$= E(b).V(\bar{x}_i)$$

$$= \frac{h^2}{2}\cdot\frac{\sigma_x^2}{n_i}$$

$$= \frac{1}{2n_i}\sigma_g^2$$

since b is orthogonal to $\bar{x}_i$ and $2E(b) = h^2 = \frac{\sigma_g^2}{\sigma_x^2}$

$\text{cov}(bA, \bar{x}) = A \ \text{cov}(b, \bar{x}_i) = 0$ since b is orthogonal to $\bar{x}_i$.
Substituting these values in (16)

$$\text{Cov}(s_i, \bar{x}_i) = \tfrac{1}{2}\text{Cov}(\bar{g}_i, g_i') + \tfrac{1}{2}\text{Cov}(g_i', \bar{\delta}_i) + \tfrac{1}{2}\text{Cov}(\bar{g}_i, \bar{\delta}_i) + \text{Cov}(\bar{g}_i, \bar{e}_i) + \text{Cov}(\bar{\delta}_i, \bar{e}_i) \quad (17)$$

It may be noted from (17) that $\text{cov}(S_i, \bar{x}_i)$ does not contain a term involving σ_g^2 as is the case with $\text{cov}(D_i, \bar{x}_i)$; neither is there a term involving σ_g^2 as is in $\text{cov}(I_i, \bar{x}_i)$. The correlation of S_i with $\bar{x}_i$ will therefore be expected to be appreciably lower than the correlation of either D_i or I_i with $\bar{x}_i$. If the conditions necessary for the covariances in (17) to vanish are satisfied S_i will not be correlated with $\bar{x}_i$. These conditions have already been discussed in Section 5.2. From these it emerges that random allotment of cows to different sires and similar conditions of feeding and management for all the daughters born are two essential prerequisites of a sound progeny testing programme.

5.5. It will be seen from the foregoing sections that the corrected index S as also the simple daughter average index $\bar{y}$ for a sire, is an estimate of $A + \frac{g'}{2}$ as against the intermediate index I which attempts to estimate $A + g'$. But this need cause no concern as $\frac{g'}{2}$ which is a constant multiple of the genotypic deviation is as good a measure as g_i'. The former is in fact the direct measure of the sire's influence on that of his daughter while the latter is a measure of the postulated genotypic value of the sire. It follows that the difference between the indices of two sires as measured by the simple daughter average index or the corrected index estimates $\frac{1}{2}(g_1' - g_2')$ as against the difference between their intermediate indices which attempts to measure $(g_1' - g_2')$. One consequence of this difference is that the intermediate index of a bull having superior genes is likely to show a greater deviation from the herd average than that of its simple daughter average or the corrected index. Similar will be the case for a bull having inferior transmitting ability. Hence the difference between the best and the worst bull used on a herd is likely to be larger when measured by intermediate index than by the other two indices. This advantage of the former is, however, offset by the likelihood of the larger bias being present. It will also be shown in the next section that the variance of the intermediate index is substantially higher than that of the other two.

Another term which occurs in the expected values of all the three indices is the herd average A. For all intra-herd comparisons this will vanish and hence no comment is required in such cases. When, however, two bulls used on two different herds are compared, the herd constants A_1 and A_2 may differ partly due to differences in management and other non-heritable causes and partly to differences in the overall genic composition of the two herds. The differences attributable to non-heritable causes should be corrected for, but no allowance need be made for the differences attributable to genetic compositions, if the correction index is used and it is assumed that there is no interaction between heridity and environment. Evidence in other countries seems to indicate that the differences in herd averages for the same breed is attributable largely to management and other environmental causes. To the extent this is true, the deviations of the indices of two sires from the respective herd averages will provide appropriate inter-herd comparisons.

Interaction between genotype and environment may vitiate comparisons, both intraherd and between herds. The possible safeguard against this in the ~~xxxxxxxfxxx~~ case of intraherd comparisons would be to allot representative lots of cows to the different bulls whereas in case of more than one herd, the bulls will have to be tested on representative lots of cows in each of the herds.

5.6 Variance of the Daughter Average Index

The sampling variance of the simple daughter average index D_i for the sire i, calculated from the records of his n_i daughters with records, is obtained as

$$V(D_i) = V(\bar{y}_i) = \frac{\sigma_y^2}{n} \qquad (18)$$

The mean square error of D_i will contain, in addition to $V(D_i)$, a variance component due to the differential production capacity of the dams whose daughters' records are used to test the sire. The additional term will be equal to the square of the bias term i.e. $(\frac{Y_i}{2})^2$. This term will be in the nature of an error variance if the bulls are allotted at random to groups of cows differing in levels of production.

The phenotypic variance σ_y^2 between the daughters of a sire will consist of a genotypic component and an environmental component. The genotypic component of σ_y^2 will be lower than σ_g^2 the genotypic variance in their dams on account of the correlation between the different daughters due to their common sire. As the correlation between half sibs (members of a family having one common parent the other parent varying) will on an average be 1/4, the genotypic variance between the daughters of a bull will be only $(1-\rho)\sigma_g^2 = (1-\frac{1}{4})\sigma_g^2 = \frac{3}{4}\sigma_g^2$

The environmental component in σ_y^2 is $E(e_{ij}^2)$

= say σ_e^2

Hence

$$V(D_i) = \frac{1}{n_i}\left(\frac{3}{4}\sigma_g^2 + \sigma_e^2\right) \qquad (19)$$

When the daughter average index is used to compare the transmitting abilities of two sires, the comparison will be rendered less precise or vitiated by the differences in the production capacities of the groups of cows allotted to the two bulls, according as the allotment is random or otherwise. Hence the expected mean square error of $(D_1 - D_2)$ is given by

$$V(D_1) + V(D_2) + \left(\frac{\gamma_1 - \gamma_2}{2}\right)^2$$
$$= \left(\frac{1}{n_1} + \frac{1}{n_2}\right)\left(\frac{3}{4}\sigma_g^2 + \sigma_e^2\right) + \frac{(\gamma_1 - \gamma_2)^2}{4} \qquad (20)$$

5.7 Variance of the Intermediate Index

The sampling variance of the intermediate index I_i of the sire i is given by

$$V(I_i) = V(2\bar{y}_i - \bar{x}_i)$$
$$= 4V(\bar{y}_i) + V(\bar{x}_i) - 4\,\mathrm{Cov}(\bar{y}_i, \bar{x}_i)$$
$$= 4V(\bar{y}_i) + V(\bar{x}_i)\{1 - 2h^2\}$$
$$= \frac{1}{n_i}\left[4\sigma_y^2 + \sigma_x^2(1 - 2h^2)\right]$$
$$= \frac{1}{n_i}\left[4\left(\frac{3}{4}\sigma_g^2 + \sigma_e^2\right) + (\sigma_g^2 + \sigma_\delta^2)(1 - 2h^2)\right]$$
$$= \frac{1}{n_i}\left[2(2 - h^2)\sigma_g^2 + 4\sigma_e^2 + (1 - 2h^2)\sigma_\delta^2\right] \qquad (21)$$

A reference to equation (9) in section 5.2 will show that in deriving the above expression it is tacitly assumed that the conditions under which the covariance terms in the eqation (9) will vanish are satisfied.

If it is assumed that the environmental variation in dams is of the same order as in their daughters, the eqation (21) simplifies to

$$V(I_i) = \frac{1}{n_i}\left[2(2 - h^2)\sigma_g^2 + (5 - 2h^2)\sigma_\delta^2\right]$$

Since $\sigma_g^2 = h^2\sigma_x^2$ and $\sigma_\delta^2 = (1 - h^2)\sigma_x^2$

$$V(I_i) = \frac{(5 - 3h^2)}{n_i}\sigma_x^2 \qquad (22)$$

From this it is seen that the standard error of the intermediate index, expressed as a percentage of the overall average, is of the order $\frac{C}{\sqrt{n}}\sqrt{5-3h^2}$, where C is the coefficient of variation and n is the number of daughters.

When two sires are compared

$$V(I_1 - I_2) = \left(\frac{1}{n_1}+\frac{1}{n_2}\right)\left[2(2-h^2)\sigma_g^2 + 4\sigma_e^2 + (1-2h^2)\sigma_\delta^2\right]$$

$$\simeq \left(\frac{1}{n_1}+\frac{1}{n_2}\right)(5-3h^2)\,\sigma_x^2 \qquad (23)$$

5.8 Variance of the Corrected Index

Sampling variance of the corrected index is given by

$$V(S_i) = V\left[\bar{y}_i - b(\bar{x}_i - A)\right]$$

$$= V(\bar{y}_i) + V(b\bar{x}_i) + V(Ab)$$
$$- 2\,\mathrm{Cov}(\bar{y}_i, b\bar{x}_i) + 2\,\mathrm{Cov}(\bar{y}_i, Ab)$$
$$- 2\,\mathrm{Cov}(b\bar{x}_i, Ab) \qquad (24)$$

Now $V(\bar{y}_i) = \frac{1}{n}\sigma_y^2$

$V(b\bar{x}_i) = \{E(b)\}^2 V(\bar{x}_i) + \{E(\bar{x}_i)\}^2 V(b)$ (covariance $\mathrm{cov}(b,\bar{x}_i)$ vanishing due to orthogonality of b and $\bar{x}_i$)

$$= \frac{1}{n_i}\left(\frac{h^2}{2}\right)^2\sigma_x^2 + \{E(\bar{x}_i)\}^2 V(b)$$

$$V(Ab) = A^2 V(b)$$

$$\mathrm{Cov}(\bar{y}_i, b\bar{x}_i) = E\{\bar{y}_i, b\bar{x}_i\} - E(\bar{y}_i)\,E(b\bar{x}_i)$$

$$= E(b)\,.\,\mathrm{Cov}(\bar{y}_i, \bar{x}_i)$$

$$= \frac{h^2}{2}\cdot\frac{h^2}{2n_i}\,\sigma_x^2$$

$$\mathrm{Cov}(\bar{y}_i, Ab) = A\,\mathrm{Cov}(\bar{y}_i, b) = 0$$

$$\mathrm{Cov}(b\bar{x}_i, Ab) = A\,\mathrm{Cov}(b\bar{x}_i, b) = A\,E(\bar{x}_i)\,V(b)$$

Substituting the values of the variances and covariances in eqation (24)

$$V(S_i) = \frac{1}{n_i}\sigma_y^2 + \frac{(h^2)^2}{4n_i}\sigma_x^2 + \{E(\bar{x}_i)\}^2 V(b)$$

$$+ A^2 V(b) - \frac{(h^2)^2}{2n_i}\sigma_x^2 - 2A\{E(\bar{x}_i)\}V(b)$$

$$= \frac{1}{n_i}\left\{\sigma_y^2 - \frac{(h^2)^2}{4}\sigma_x^2\right\}$$

$$+ V(b)\left[\{E(\bar{x}_i)\}^2 + A^2 - 2A \cdot E(\bar{x}_i)\right]$$

$$= \frac{1}{n_i}\left\{\sigma_y^2 - \frac{(h^2)^2}{4}\sigma_x^2\right\} + \{E(\bar{x}_i) - A\}^2 V(b) \quad (25)$$

When b is estimated from data pertaining to k bulls, the i-th bull having n_i daughter-dam pairs,

$$V(b) = \frac{\left[\sigma_y^2 - \frac{(h^2)^2}{4}\sigma_x^2\right]}{n-k} \qquad \text{where} \quad n = \sum_{i=1}^{k} n_i$$

Substituting in (25)

$$V(S_i) = \left[\sigma_y^2 - \frac{(h^2)^2}{4}\sigma_x^2\right]\left[\frac{1}{n_i} + \frac{\{E(\bar{x}_i) - A\}^2}{(n-k)\sigma_x^2}\right]$$

$$= \sigma_x^2\left[\frac{\sigma_y^2}{\sigma_x^2} - \frac{(h^2)^2}{4}\right]\left[\frac{1}{n_i} + \frac{\{E(\bar{x}_i) - A\}^2}{(n-k)\sigma_x^2}\right] \quad (26)$$

If the environmental variation for the daughters and the dams may be taken to be of the same order,

$$\sigma_e^2 = \sigma_\delta^2$$

and $\sigma_y^2 = \left(\frac{3}{4}\sigma_g^2 + \sigma_e^2\right) = \sigma_x^2\left(1 - \frac{h^2}{4}\right)$

Hence $V(S_i) = \left[1 - \frac{(h^2)}{4} - \frac{(h^2)^2}{4}\right]\left[\frac{1}{n_i} + \frac{\{E(\bar{x}_i) - A\}^2}{(n-k)\sigma_x^2}\right] \quad (26a)$

The term $\frac{\{E(\bar{x}_i) - A\}^2}{(n-k)\sigma_x^2}$ may be neglected compared to $\frac{1}{n_i}$ if $E(\bar{x}_i)$ is quite close to the herd average A

In this case

$$V(S_i) = \frac{1}{n_i}\left[1 - \frac{(h^2)}{4} - \frac{(h^2)^2}{4}\right]\sigma_x^2 \qquad (26b)$$

From this it is seen that the standard error of the corrected index expressed as a percentage of the herd average, is of order

$$\frac{c}{\sqrt{n}}\sqrt{1 - \frac{(h^2)}{4} - \frac{(h^2)^2}{4}}$$

where C and N are respectively the coefficient of variation and the number of daughter-dam pairs.

When two sires are compared $V(S_1 - S_2)$ can be shown to be

$$= \left\{\sigma_y^2 - \frac{(h^2)^2}{4}\sigma_x^2\right\}\left[\frac{1}{n_1} + \frac{1}{n_2} + \frac{\{E(\bar{x}_1) - E(\bar{x}_2)\}^2}{(n-k)\sigma_x^2}\right] \qquad (27)$$

5.9 Relative efficiencies of the indices

The relative efficiency of the corrected index compared to the simple daughter average index from eqations (18) and (26) is

$$= \frac{\frac{1}{n_i}\sigma_y^2}{\left\{\sigma_y^2 - \frac{(h^2)^2}{4}\sigma_x^2\right\}\left[\frac{1}{n_i} + \frac{\{E(\bar{x}_i) - A\}^2}{(n-k)\sigma_x^2}\right]}$$

$$= \frac{\frac{1}{n_i}\left(\frac{3}{4}\sigma_g^2 + \sigma_e^2\right)}{\left\{\left(\frac{3}{4}\sigma_g^2 + \sigma_e^2\right) - \frac{(h^2)^2}{4}(\sigma_g^2 + \sigma_e^2)\right\}\left[\frac{1}{n_i} + \frac{(r_i - A)^2}{(n-k)(\sigma_g^2 + \sigma_e^2)}\right]} \qquad (28)$$

Assuming that $\sigma_e^2 = \sigma_\delta^2$ and that the ter m $\frac{(r_i - A)^2}{(n-k)\sigma_x^2}$ is negligible compared to $\frac{1}{n_i}$

$$\frac{V(D_i)}{V(S_i)} = \frac{\left(1 - \frac{h^2}{4}\right)}{\left[1 - \frac{h^2}{4} - \frac{(h^2)^2}{4}\right]} \qquad (28a)$$

$$= \left(1-\frac{h^2}{4}\right)\left[1+\left\{\frac{h^2}{4}+\frac{(h^2)^2}{4}\right\}+\left\{\frac{h^2}{4}+\frac{(h^2)^2}{4}\right\}^2+\cdots\right]$$

$$= \left(1-\frac{h^2}{4}\right)\left[1+\frac{h^2}{4}+\frac{5}{16}(h^2)^2+\cdots\cdots\right]$$

$$= 1+\frac{(h^2)^2}{4} + \text{higher powers of } h^2 \qquad (28\,b)$$

Hence gain in efficiency $= \frac{(h^2)^2}{4}$ to the order of $(h^2)^2$

Relative efficiency of the Corrected index compared to intermediate index is

$$\frac{V(I_i)}{V(S_i)} \doteq \frac{\frac{1}{n_i}\left[4\sigma_y^2+\sigma_x^2(1-2h^2)\right]}{\left\{\sigma_g^2-\frac{(h^2)^2}{4}\sigma_x^2\right\}\left[\frac{1}{n_i}+\frac{\{E(\bar{x}_i)-A\}^2}{(n-k)\sigma_x^2}\right]} \qquad (29)$$

$$\simeq \frac{\left[4\left(1-\frac{h^2}{4}\right)+(1-2h^2)\right]}{\left\{1-\frac{h^2}{4}-\frac{(h^2)^2}{4}\right\}}$$

$$= \frac{(5-3h^2)}{\left\{1-\frac{h^2}{4}-\frac{(h^2)^2}{4}\right\}} \qquad (29\,a)$$

$$= (5-3h^2)\left\{1-\frac{h^2}{4}-\frac{(h^2)^2}{4}\right\}^{-1}$$

$$= (5-3h^2)\left[1+\left\{\frac{h^2}{4}+\frac{(h^2)^2}{4}\right\}+\left\{\frac{h^2}{4}+\frac{(h^2)^2}{4}\right\}^2+\cdots\right]$$

$$= (5-3h^2)\left[1+\frac{1}{4}h^2+\frac{5}{16}(h^2)^2+\cdots\right]$$

$$= 5+\left(\frac{5}{4}-3\right)h^2+\left(\frac{25}{16}-\frac{3}{4}\right)(h^2)^2+\cdots$$

$$= 5-\frac{7}{4}h^2+\frac{13}{16}(h^2)^2+\cdots \qquad (29\,b)$$

Hence gain in efficiency $= 5-\frac{7}{4}h^2+\frac{13}{16}(h^2)^2$ to the order of $(h^2)^2$.

The gain in efficiency of the corrected index relative to the simple daughter average index and to the intermediate index for values of the heritability coefficient ranging from .1 to 1.0 are indicated in the table below.

Percentage gain in efficiency of the corrected index.

H^2	Over simple daughter average index	Over Intermediate Index
.1	Less than .5	383
.2	1	368
.3	2	355
.4	4	343
.5	6	332
.6	9	324
.7	12	317
.8	16	312
.9	20	308
1.0	25	306

The gain in efficiency of the corrected index over the intermediate index is over 330% for a heritability coefficient of 0.5 or less. The gain of the former index over the daughter average index is very small (6% or less) for h^2 less than 0.5 when representative sets of cows are allotted to the different bulls under test. It must be remembered, however, that the utility of the corrected index lies in eliminating the bias due to the unequal production level of dams from which the simple daughter average index suffers.

5. Empirical verification

In order to see how far the expectation of the superiority of the corrected index over the other two indices is realised in practice, data relating to six Indian dairy herds were analysed. The data studied consisted of 990 daughter-dam pairs from 69 sires. Details of the computational procedure adopted for calculating the estimate and the standard error of the corrected index, illustrated on one of the herds, is given in the Appendix.

The three indices for each of the bulls tested in the different herds is given in Tables 1 to 6. It will be seen from columns 7 and 9 of the tables that the standard errors for the corrected and the intermediate indices are, generally speaking, in the ratio of 1:2 as was expected. A comparison of the standard errors for the corrected and the simple average indices indicates that the gain achieved due to the reduction in the standard errors is negligible.

A comparison of the efficiencies of the estimation can be made directly by a comparison of the mean square errors for the three indices. It will be seen from Table 7 that the mean squares for the intermediate and the simple average indices, when averaged over all the six herds, were about 437 and 103 per cent respectively of the mean square for the corrected index. The standard error of the corrected index, when averaged over all sires, was only 45.4 per cent of the standard error for the intermediate index and 97.8 per cent of the error for the simple daughter average index.

Comparison of different sire indices – Kankrej bulls, Livestock Research station, Hosur.

Bull No.	Daughter dam pairs.	Dams' Average	Simple Daughter Average		Intermediate		Corrected Daughter Average@		Rank according to		
			Index	S.E.	Index	S.E.	Index	S.E.	(4)	(6)	(8)
1	2	3	4	5	6	7	8	9	10	11	12
307	54	1673	1176	101	679	168	1098	90	10	15	16
306	45	1818	1245	94	672	170	1155	88	12	16	13
35	34	14[illegible]2	1046	91	1650	207	1534	96	9	7	7
90	30	1661	1746	123	1835	217	1675	111	3	5	4
269	28	1645	1249	87	853	190	1179	91	11	12	12
231	26	1518	1182	77	846	164	1148	76	15	13	14
85	20	129[illegible]	1654	90	2016	212	1685	96	6	3	3
391	1[illegible]	1408	1185	59	962	161	1183	66	14	10	11
104	11	2421	1706	180	991	571	1414	225	5	9	9
50	9	1673	1717	212	1761	362	1639	193	4	6	6
33	9	1203	1608	115	2013	174	1665	89	7	4	5
34	8	1806	1585	115	1364	347	1468	139	8	8	8
132	8	17[illegible]9	1340	148	951	394	1246	166	10	11	10
5[illegible]4	7	1894	1073	170	252	391	931	172	17	17	17
39	7	2045	2360	482	2674	899	2174	466	1	1	1
47	7	1614	1186	115	758	320	1124	136	13	14	15
119	6	1873	2135	289	2397	547	1999	27[illegible]	2	2	2

@ Corresponding to the herd average of 1400 lbs.

-Note-

Comparison of different sire indices in Tharparkar bulls, Government Cattle Farm, Patna.

Bull No.	Daughter dam pairs.	Dams' average	Simple Daughter Average		Intermediate		Corrected Daughter average@		Rank according to		
			Index	S.E.	Index	S.E.	Index	S.E.	(4)	(6)	(8)
1	2	3	4	5	6	7	8	9	10	11	12
39	45	3340	2704	172	2068	416	2679	176	10	11	10
26	29	3297	3128	193	2959	423	3110	193	4	5	4
599	26	3441	3160	261	2879	542	3128	263	3	6	3
24	24	3948	3020	277	2092	511	2951	283	5	10	6
1461	21	3594	3345	315	3096	648	3342	319	2	3	2
77	20	3310	2780	268	2250	438	2758	261	9	8	9
20	14	3272	2876	211	2480	381	2856	206	8	7	8
16/0	14	3783	3639	374	3495	864	3582	385	1	1	1
885	13	3484	2160	161	836	541	2125	190	13	13	13
18	10	3101	2673	172	2245	351	2666	168	11	9	11
16	8	2694	2891	256	3088	649	2913	258	7	4	7
27	6	3700	2605	565	1510	1550	2554	539	12	12	12
1/0	6	2625	3002	721	3379	1394	3029	654	6	2	5

@ Corresponding to the herd average of 3000 lbs.

Note

Table 7

[illegible] Sindhi Bulls, Live stock Research St[illegible], [illegible].

Bull No.	Daughter-dam pairs.	Dams' average	Simple Daughter average Index	S.E.	Inter[illegible] Index	S.E.	Corrected Daughter average@ Index	S.E.	Ran[illegible] (4)	(6)	( )
1	2	3	4	5	6	7	8	9	10	11	12
255	27	4019	2974	239	1229	497	2917	236	13	13	13
8	25	3651	4363	272	5075	506	4372	258	3	3	2
144	23	4118	3192	315	2366	713	3117	323	12	11	12
236	18	3860	3923	475	3986	995	3895	477	7	6	7
38	12	3664	4023	226	4382	396	4029	212	6	4	6
65	12	4059	3268	269	2495	641	3203	282	11	10	10
139	11	4827	4261	362	3695	509	4060	303	4	7	5
282	11	4320	3[illegible]78	481	2236	1076	3167	490	10	12	11
124(0)	11	3364	3283	495	3202	953	3343	481	9	8	9
98	10	4499	3813	360	3127	557	3070	310	8	9	8
86	9	4298	5083	850	5868	1698	4976	841	1	1	1
238	7	4646	4470	426	4294	1096	4301	401	2	5	3
136(0)	6	2098	4163	318	5328	644	4286	317	5	2	4

* Note *

@ Corresponding to the herd average of 3700 lbs.

Comp[illegible] [illegible] [illegible] 5 -
Sindhi Bulls, Indian Dairy Research
Institute, Bangalore.

Name of Bull	Daughter-dam pairs	Dams' Average	Simple Daughter Average Index	Simple Daughter Average S.E.	Intermediate Index	Intermediate S.E.	Corrected Daughter Average @ Index	Corrected Daughter Average @ S.E.	Rank according to (4)	Rank according to (6)	Rank according to (8)
1	2	3	4	5	6	7	8	9	10	11	12
Erab	13	[illegible]	[illegible]	308	2035	617	2577	302	8	10	9
Ziman	1[illegible]	[illegible]	[illegible]	299	2328	575	2587	279	10	8	8
Unique	1[illegible]	[illegible]541	[illegible]	318	4025	729	3312	324	3	2	2
Erio	17	4117	[illegible]1	430	1445	853	2519	424	7	11	10
M. Gulam	12	5131	[illegible]	368	3255	767	2798	362	9	5	6
Sulaiman	10	[illegible]	[illegible]43	523	2403	723	2735	452	6	7	7
Tazman	10	[illegible]	31[illegible]	266	3496	507	3160	242	4	4	4
Sikander	9	[illegible]	[illegible]6	457	2550	1152	2425	492	11	9	11
Lavier	9	47[illegible]	[illegible]	578	3647	1499	3104	628	5	3	5
H. Raj	7	3777	1810	573	847	973	1797	525	12	12	12
Warrior	7	38[illegible]7	[illegible]60	442	2[illegible]29	1061	3163	462	2	6	3
Victory	6	3679	3010	201	4959	651	3823	536	1	1	1

@ Corresponding to the herd average of 2700 lbs.

-N.R.

Comparison of different Sire Indices –
Hariana Bulls, Government Cattle Farm, Hissar.

Bull No.	Daughter Dam pairs	Dams' average	Simple Daughter Average Index	Simple Daughter Average S.E.	Intermediate Index	Intermediate S.E.	Corrected Daughter Average@ Index	Corrected Daughter Average@ S.E.	Rank according to (4)	Rank according to (6)	Rank according to (8)
1	2	3	4	5	6	7	8	9	10	11	12
I	[illegible]	1334	1329	[illegible]	1324	222	1319	101	8	9	9
II	[illegible]	[illegible]	[illegible]	151	910	274	1167	139	10	10	10
III	10	[illegible]	[illegible]	[illegible]	1533	440	1523	210	2	7	3
IV	9	[illegible]	[illegible]	139	1542	344	[illegible]	153	7	6	7
V	[illegible]	114[illegible]	[illegible]	[illegible]	1731	213	1483	101	5	2	4
VI	7	[illegible]	1717	[illegible]	2241	385	1440	182	1	1	1
VII	6	[illegible]	[illegible]	[illegible]	1632	84	1539	[illegible]	6	3	6
VIII	7	[illegible]	[illegible]	230	1597	371	1470	200	3	5	2
IX	11	[illegible]	[illegible]	165	1607	364	1340	[illegible]	4	4	5
X	5	[illegible]	[illegible]	[illegible]	1454	378	1748	182	9	8	8

@ Corresponding to the herd average of 1300 lbs.

30/3/57

Table 6

Comparison of different Sire Indices -
Gir Bulls, Indian Dairy Research Institute,
Bangalore.

Name of Bull	Daughter-dam pairs.	Dams' Average	Simple Daughter Average Index	Simple Daughter Average S.E.	Intermediate Index	Intermediate S.E.	Corrected Daughter Average@ Index	Corrected Daughter Average@ S.E.	Rank according to (4)	Rank according to (6)	Rank according to (8)
1	2	3	4	5	6	7	8	9	10	11	12
Windfall	10	3368	3123	222	2878	280	3009	162	1	2	2
Yakub Khan	8	2102	2163	393	2134	631	2414	383	4	3	3
Ranji	5	2598	3057	444	3516	1063	31[illegible]	499	2	1	1
Wonderful	5	3189	2343	614	1497	14[illegible]	2284	672	3	4	4

@ Corresponding to the herd average of 3000 lbs.

* [illegible] *
30/3/57

Table - 7

Pooled mean squares errors for the different indices (10 lb. units).

Herd	Degrees of freedom	Mean Square for		
		Simple daughter average	Intermediate	Corrected
1. Kangayam (Hosur)	304	3536	12166	3173
2. Tharparkar (Patna)	223	13936	63850	13938
3. Sindhi (Hosur)	174	21177	88410	20855
4. Sindhi (Bangalore)	127	18467	80645	17716
5. Hariana (Hissar)	69	2558	9425	2189
6. Gir (Bangalore)	24	10248	38857	9806
Pooled.	921	11547	48998	11219

Table 8

Correlation of the different indices with the average of the dams

Herd	Number of sires tested	Simple daughter average index		Intermediate index		Corrected index	
		Directly estimated	Expected	Directly estimated	Expected	Directly estimated	Expected
Kangayam, Hosur	17	0.30	0.32	- 0.10	- 0.26	0.08	Zero
Tharparkar, Patna	13	0.18	0.06	- 0.35	- 0.36	0.11	"
Sindhi, Hosur	13	0.15	0.14	- 0.26	- 0.24	0.005	"
Sindhi, Bangalore	12	0.07	0.22	- 0.44	- 0.36	- 0.16	"
Hariana, Hissar	10	0.42	0.41	- 0.36	- 0.29	- 0.04	"
Combined	65	0.22	0.23	- 0.28*	- 0.30	- 0.04	"

* Significant at 5 per cent level.

It was shown in section 4 that under the assumptions required for the covariance terms in equation (9) to vanish, the daughter average index of a bull and its intermediate index will respectively be positively and negatively correlated with the average yield of dams. Under the same assumptions it was concluded that correlation between corrected index and dams' average is absent . An empirical verification of these results was undertaken by working out correlation between each of the three indices and the average yield of the dams which were the mates of the different sires. The values of the correlations obtained for the different herds studied are presented in Table 8. The correlation between the simple daughter average index and the dams' average was positive in all the herds. Similarly the intermediate index was uniformly negatively correlated with the average yield of the dams. (Similar results were also observed to hold for the Gir herd studied, the actual values of the correlation for which herd is not presented in the table as the number of bulls tested were only four). The correlation between the corrected index and the dams average was positive for three herds and negative for the other two.

The expected values of these correlation coefficients using the heritability and the variance estimates (viz $h^2, \sigma_x^2, \sigma_y^2$) were also obtained and are presented in the same table for comparison. The agreement appears to be satisfactory on the whole, the pooled values almost coinciding with their expectations. The few discrepancies noted in the comparisons within herds are to be ascribed to the small number of the bulls tested.

T a b l e 9

Rank correlation between different sire indices

H e r d	Number of sires	Simple daughter average and intermediate	Corrected and simple daughter average	Corrected and intermediate
Kangayam (Hosur)	17	0.87	0.92	0.98
Tharparkar(Patna)	13	0.81	0.997	0.86
Sindhi (Hosur)	13	0.89	0.98~~4~~	0.94
Sindhi (Bangalore)	12	0.70	0.91	0.93
Hariana (Hissar)	10	0.70	0.96	0.78

The relative rankings of the bulls in each breed according to the three indices are given in columns (10), (11) and (12) of tables ~~III to VIII~~ 1 to 6. It may be seen that for all the herds studied, the bull ranked first according to the corrected index retains the same rank according to the intermediate index or to the ordinary daughter average index. The only exception was the Gir herd, where the best bull would have been rated as the second best if the ordinary daughter average index has been used. For all the six ~~breeds~~ herds, the worst bull would have been rated as such according to any of the three indices. The rankings for other bulls also are in close agreement. The values of the rank correlation of the corrected index with the intermediate index and the simple daughter average index are given in Table 9. The rank correlations are uniformly high for all the herds showing thereby that the use of the simple daughter average index will seldom lead to materially different conclusions, if the object is merely to order the bulls according to their relative merits, provided that the average production levels of the dams mated to different bulls do not differ to a great extent. Whenever such variations are present, or when the estimates of the sire indices are desired with a greater precision, the corrected index is to be preferred.

7. Discussion

In the model discussed in Section 5, it was assumed that each cow had only one daughter with records. This simplifying assumption will not generally hold for actual data collected from breeding farms. In almost every farm it is a common practice to retain high yielding cows for a larger number of lactations with the consequence that these cows will have more daughters than the poor yielders.

If the mating system followed is such that each daughter from a cow is raised through a different sire, the procedure for the calculation of any of the three indices considered, needs little alteration. The average level of production of the mates will, however, be increased by the selection exercised on them from lactation to lactation. This will raise the herd average; but will not vitiate the comparison of two sires if they are used on cows with contemporary records and the cows allotted to each sire is a representative sample of the herd. In case the groups of cows assigned to each bull are different, the situation will remain essentially the same as has been discussed earlier. The worth of a sire used on cows having relatively higher average is likely to be over-estimated if the simple daughter average index is used and under-estimated if the intermediate index is employed, owing to the positive and the negative correlation respectively of these indices with the dam's yield. The corrected index, being uncorrelated with the dam's yield, is not likely to introduce such a bias.

When a number of cows are retained for a large number of calvings a procedure that can be followed to eliminate the influence of the unequal production capacity of the dams is to obtain a daughter through the mating of each cow with each sire. The scope of this method of sire testing, called diallel crossing, is limited in practice by the undeterminable nature of the sex of the calf born and the relatively few calves that can be raised from a cow. Further this method also requires that every cow used for test should be retained for a large number of lactations. This method may seem to be useful when only a few sires, say two or three, are required to be tested on a large herd. Even in this case its value is affected by the longer time interval required for the completion of the tests.

In situations where selection is practised on cows in successive lactations and a cow is allotted ~~to~~ more than once to the same bull, bias in the estimation of the sire index is likely to arise from two sources. One source is that cows retained for more lactations are likely to contribute a larger proportion of the mates of a bull, the effect of which is analogous to the case where the dams level of production is higher than the herd average. Allotment of such cows to different bulls is a second source of bias which may be eliminated if the bull to be used for service is determined independently for each mating.

With two or more daughters of a cow sired by the same bull, two practices have been widely used in sire testing. One is to repeat the dam's record with each daughter's record. The other is to average the records of all the daughters forming a full-sib family and consider this average along with the dam's record as constituting a daughter-dam pair. The former practice would be valid if the correlation between the daughters constituting a full sib family were zero while the latter would be appropriate if this correlation is perfect. Obviously the actual situation in most of the breeding material is intermediate to these two extreme conditions, although usually nearer to the former. The procedure of repeating the dam's record with each daughter's record is, therefore, to be preferred over the other.

As has been stated earlier, the most efficient estimate of the heritability coefficient is obtained by doubling the intrasire regression of the daughter's performance on that of its dam. The influence of the varying number of offspring per parent on the regression of the offspring value on that of its parent, was considered by Kempthorne and Tandon (1953). The procedure recommended by them was a system of weighting which will give an unbiassed estimate of the regression with minimum sampling variance. It may be briefly summarised as follows:

I Guess $T = \frac{\rho}{(1-\rho)}$ where ρ is the correlation between the full sibs and call the guessed value τ

II The estimate of heritability β is given by

$$\hat{\beta} = \frac{\sum_j \omega_j (x_j - \tilde{x}) y_j}{\sum \omega_j (x_j - \tilde{x})}$$

where $\omega_j = \frac{n_j}{1+n_j\tau}$, n_j being the number of daughters of the jth cow from a bull and $\tilde{x} = \frac{\sum_j \omega_j x_j}{\sum_j \omega_j}$

III Estimate ρ_1 by using the intra-sire mean squares between daughters, within and between cows. The expectation of the mean square within cows is $\sigma_p^2(1-\rho_1)$ and of the mean square between cows is

$$\left[\sigma_p^2(1-\rho_1) + \frac{1}{k-1}\left(\sum_j n_j - \frac{\sum_j n_j^2}{\sum_j n_j}\right)\sigma_p^2\rho_1\right]$$

where σ_p^2 is the phenotypic variance and

k is the number of cows, so that one can estimate ρ_1 by equating observed to expected mean squares.

IV. Using the estimates $\hat{\rho}_1$ and $\hat{\beta}$ obtain

$$\hat{T} = \frac{\hat{\rho}_1 - \hat{\beta}^2}{1-\hat{\rho}_1}$$

V The estimated variance of $\hat{\beta}$ is then

$$= \left[\sum_j n_j \frac{1+n_j\hat{T}}{(1+n_j\tau)^2} \frac{(x_j-\tilde{x})^2}{\sum_j \omega_j (x_j-\tilde{x})^2}\right] \times \text{M.S. between daughters within cows}$$

The success of the use of such a procedure obviously depends on the closeness of the value of T_f to τ it is also somewhat combursome. In most situations dealing with quantitative character, the h^2 being rather small, the simple procedure of repeating the dam's record with each daughter's record may provide a satisfactory approximation and is to be preferred to unweighted regression of means of daughters on dams.

For sire évaluation and other breeding studies, it is generally the practice in foreign countries and also in some farms in India, to take the yield in the first 300 (or 305) days or the complete lactation yield when the period is less than 300 days. When such data are not available (as was the case in respect of the herds taken for illustration excepting the Hariana), Sukhatme (1944) suggested that correction for the inequality in lactation period may be made by using the regression technique with lactation period as the concomitant variate. This method, however, inflates the indices of those sires whose daughters have shown poor performances since poor performers generally have shorter lactation periods. Raising the yield of daughters which have ceased to give milk much earlier than 300 days, does not appear to be justified as the shorter lactation length cannot be ascribed wholly to random environmental causes. The lower lactation lengths are at least in part due to poorer genotypes. The actual lactation yield rather than the yield adjusted for 300 days should, therefore be considered as reflecting the milk potentiality of the progeny. Cases where shorter lactation periods are due to the result of known abnormalities such as death of a calf or diseased condition of the cow, should be omitted rather than corrected. For cows having a lactation period longer than 300 days also the yield corrected to 300 days by using regression technique does not appear to be a suitable substitute to the actual yield obtained during that period, as the regression results generally in an over-correction. This is on account of the fact that the yield in the tail end of a lactation is generally lower than the average yield per day over the entire lactation period. It seems, therefore desirable that the yield in the first 300 (or 305) days of a lactation should be taken for the lactations longer than this period. A good reason for this practice is that in dairy farms, it is desirable to aim at annual calvings and to provide for a dry period of about two months to help the cow maintain her health. For lactations completed in less than 300 days, the complete lactation record seems to be the appropriate one. Wherever data on the yield in the first 300 days are not available, it is preferable to carry out the studies on the unadjusted yield rather than on the yield corrected to 300 days using the regression technique

Another factor influencing the lactation yield of a cow is the order of lactation. The effect of this factor can be eliminated by confining the study to the first lactation records only. The first lactation records are preferable to the later ones as they will be available earlier and will be influenced to a lesser extent by selection. The extent of gain that can be achieved by using the later lactation records in addition, requires further investigation.

An important consideration in planning a systematic breeding programme providing for sire evaluation is the number of daughter-dam pairs required to prove a sire. An answer to this question depends on how sure of his proof one wants to be and on the order of variability among the daughters' yields after correcting for the inequalities in the dams' performances. With the conventional five per cent level of significance and a coefficient of variation of the order of forty per cent for lactation record observed in the case of the six herds already referred to, the superiority of a bull whose corrected index is twenty per cent higher than the herd average can be detected with twelve daughter-dam pairs. With the same number of pairs, it will be possible to distinguish between bulls whose corrected indices differ by more than 32 per cent at the same level of significance. Sukhatme (1944) suggested that in dealing with breeding studies the ten per cent level of significance may be used as an aid to possible retention of superior breeding material which is difficult to select with greater certainty. If this level of significance is adopted, a difference of the order of about 28 per cent or higher will be revealed as significant. The corrected index for a sire calculated on twelve pairs of records is expected to be determined with a standard error of the order of eleven per cent.

The results indicate that, while there is no harm in using as low a number of pairs as five or six, for getting the first indication of the breeding worth of a sire, and discarding inferior bulls having lower indices than the herd average, the final selection of a bull as proven for extensive use, and for selection of male breeding stock for further propagation should be based on a test carried out on the records of twelve or more daughter-dam pairs. In order to provide records of twelve daughters, about thirty dams may have to be mated to a sire so as to make allowance for sex ratio, mortality of calves etc.

8. Summary

In any scientific breeding programme for improvement of dairy cattle greater importance is attached to sire selection on account of the larger number of progenies that can be raised by a bull in his life time compared to a cow.

As milk yield is a sex limited character, sire testing has to be based on the yields of the relatives of the bull. Indices based on such records proposed by different authors have been briefly reviewed and the merits and limitations of the two common indices widely in use, namely the Simple daughter average index and the Intermediate index are examined and a new index called the corrected index, comparatively free from some of these limitations, is developed.

The proposed index is obtained from the formula $S = \bar{y} - b(\bar{x} - A)$ where $\bar{y}$ is the simple average of all the daughters of the given sire, $\bar{x}$ the average of the dams of these daughters, A the herd average and b the intrasire regression of daughter's performance on that of its dam.

By setting up a simple genetic model, the expectations and the variances of the two common indices mentioned earlier, along with that of the proposed index, are derived. The nature of the bias involved in their use is examined and the relative efficiencies of the three indices assessed. It is shown that the gain in efficiency of the corrected index over the intermediate index is 350 per cent or more for a heritability of 0.5 or less. When the value of the heritability is more than 0.5 the corrected index proved to be more efficient although to a lesser degree. In so far as the simple daughter average index is concerned, it is subject to bias on account of the unequal levels of dams' production, whereas the corrected index which is free from the bias has an efficiency in no case lower than that of former.

The superiority of the corrected index over the other two is also empirically verified with the help of data pertaining to sixty nine bulls from six herds.

Situations arising due to more than one daughter with records available for each cow is discussed. Methods for adjusting for the inequality of lactation length and the order of lactation have also been discussed. The minimum number of daughter-dam pairs required to test a sire by the use of the proposed index has been examined.

ACKNOWLEDGMENT.

I am greatly indebted to Shri V.N. Amble, Statistician (Animal Husbandry), I.C.A.R. and Dr. V.G. Panse, Statistical Adviser to the Council for their guidance and help in the preparation of this thesis.

With a view to obtaining the empirical verification on the records from as many herds as possible, a study of all the six herds for which records were available at the Council was undertaken. This involved huge volume of computation. The help rendered by Messrs P.N. Soni and B.B. Nayar, members of the Statistical staff of the Council, in working out some of the calculations pertaining to the sums, sums of squares, etc., is acknowledged with thanks.

R E F E R E N C E S

1. Berge, S (1944) — Nord. Jordler Forskn, 18.

2. Hansson, N (1913). — Fiihl. Landun. Zeit, 62.

3. Kempthorne O and Tandon, O B. (1953) — Biometrics 9.

4. Krüger, L. (1953) — Züchtungskunde 13.

5. Lush, J.L (1949) — Animal Breeding Plans, Iowa State College Press, Ames, Iowa.

6. Rice, V.A. (1944) — Jour. Dairy Science 27.

7. Sukhatme, P.V. (1944) — Ind Jour of Vet +Sci.14.

8. Von Patow, C (1930) — Zücht. B. 17.

9. Wright, S (1931) — Proc. Amer. Soc.An. Prod

10. Yapp, W.W. (1925) — Proc Scot, Cattle Breed Conf.

Table I

First Lactation Yields (in 10 lbs in 300 days or less) of the ~~dams~~ cows mated to different sires and their daughters
(Hariana herd, Government Cattle Farm, Hissar)

Sire I		Sire II		Sire III		Sire IV		Sire V		Sire VI		Sire VII		Sire VIII		Sire IX		Sire X	
Dam's Yield	Daughters Yield	Dam's Yield	Daughters Yield	Dam's Yield	Daughters Yield	Dam's Yield	Daughters Yield	Dam's Yield	Daughters Yield	Dam's Yield	Daughters Yield	Dam's Yield	Daughters Yield	Dam's Yield	Daughters Yield	Dam's Yield	Daughters Yield	Dam's Yield	Daughters Yield
176	134	127	121	108	261	200	141	68	93	186	129	235	202	43	128	47	123	94	114
222	118	231	142	177	219	108	82	74	99	67	158	28	77	168	161	152	92	180	116
"	174	207	196	"	135	91	156	87	151	208	283	73	117	"	107	229	220	138	177
82	136	109	139	228	211	107	158	232	183	"	160	"	125	219	230	62	60	87	166
32	66	154	85	105	119	179	86	141	164	68	190	268	2[illegible]7	"	139	62	172	54	71
65	110	"	162	204	190	57	152	"	171	73	151	26	93	238	272	187	302		
66	131	"	84	"	199	56	176	123	1[illegible]1	25	131			161	130	"	223		
201	194	142	74	"	120	164	104	50	168							19	178		
				203	176	78	79									209	94		
				"	43											86	135		
																273	14		

* N[illegible]*
28/3/57

Note: The symbol " indicates that the dam is the same as for the previous pair.

APPENDIX

Computational Procedure

The computation of the new index from actual data will be illustrated with the help of the records pertaining to the Hariana herd maintained at the Government cattle Farm, Hissar. The first lactation yield over 300 days or less, measured in units of ten pounds, are analysed. These yields in respect of each of the ten sires tested are given in Table I.

The quantities that are required to be calculated are (a) the herd average (b) the intra-sire regression coefficient (c) the corrected index and (d) the standard error of the index.

The herd average required for the calculation of the corrected index is obtained as the average of the dam's yields without repeating the dams having more than one daughter. This average worked out to 130.8 units in the present case. The average for the herd will be taken as 1300 lbs in the nearest round figure.

The steps in the calculation of intra-sire regression are shown in columns (3) to (13) in Table II. As an illustration the procedure for obtaining these figures for one bull, viz. sire I, is explained, below:-

Total for dams mated, repeating each dam's records as many times as the number of daughters from the same sire

= 176 +222+222+82 + 32 + 66 +66 + 201

= 1067 ... ... (Col.3)

Total for daughters

= 134 + 118 + 174 + 136 + 66 +110 +131 + 194

= 1063 ... (Col. 4)

-/-

Table II

Calculation of Corrected Daughter Average Indecies and their Standard Errors (10 lbs. unit)
(Hariana herd, Government Cattle Farm, Hissar).

Sire No.	Daughter dam pairs	Total for dams	Total for daughters	Sum of squares for dams			Sum of squares for daughters			Sum of products		
				Crude	C.T.	Corrected	Crude	C.T.	Corrected	Crude	C.T.	Corrected
1	2	3	4	5	6	7	8	9	10	11	12	13
I	8	1067	1063	186405	142311	44094	151905	141246	10659	156572	141778	14794
II	8	1278	1003	215532	204160	11372	138543	125751	12792	165374	1602[illegible]9	5145
III	10	1813	1673	344597	328697	15900	315915	279893	36022	299742	303315	-3573
IV	9	1040	1214	142800	120178	22622	177478	163755	13723	138410	140284	-1874
V	8	916	1150	128684	104882	24002	173742	165312	8430	13[illegible]761	131675	8086
VI	7	835	1202	136191	99604	36587	223356	206401	16955	153942	143381	10561
VII	6	703	841	139167	82368	56799	136225	117880	18345	130546	98537	32009
VIII	7	1216	1167	236784	211236	25548	216859	194556	22303	217005	202725	14280
IX	11	1515	1641	281695	208657	73038	274619	244807	29812	237678	2[illegible]6010	11668
X	5	561	644	73709	62944	10765	90378	82947	7431	752[illegible]5	72257	2969
Total	79	10944	11598	1885764	1565037	320727	1899020	1722548	176472	1714256	1620191	94065

* N....*
28/3/57

[illegible]

Sire No.	Av. for dams (3)-(2)	Dams Av. - Herd Av. (14)-130	(15)xb	Av. for daughters (4)-(9)	Corrected Index (17)-(16)	(9)xb^2	(13)x2b	Residual S.S. (10)+(19)-20	n(n-1)	Variance of index (21)-(22)	S.E. of Index $\sqrt{(23)}$
~~14~~	~~15~~ 14	~~16~~ 15	~~17~~ 16	~~18~~ 17	~~19~~ 18	~~20~~ 19	~~21~~ 20	~~22~~ 21	~~23~~ 22	~~24~~ 23	~~25~~ 24
I	135.4	+5.4	+1.0	132.9	131.9	3793	8676	5774	56	103	10.1
II	159.8	+29.8	+8.7	125.4	116.7	978	3018	10752	56	192	13.9
III	181.3	+51.3	+15.0	167.3	152.3	1366	32086	39486	90	439	21.0
IV	115.6	-14.4	-4.2	134.9	139.1	1946	-1099	16768	72	233	~~15.3~~ 15.3
V	114.5	-15.5	-[illegible].5	143.8	148.3	2065	4743	5752	56	103	10.1
VI	119.3	-10.7	-3.1	171.7	174.8	3147	6196	13[illegible]07	42	331	18.2
VII	117.2	-12.8	-3.8	140.2	144.0	4866	18776	4458	30	148	12.2
VIII	173.7	+43.7	+12.6	166.7	153.9	2198	8377	16124	42	364	20.0
IX	137.7	+7.7	+2.2	149.2	147.0	6283	6644	29251	110	266	16.3
X	112.2	-17.8	-5.2	128.8	134.0	926	1745	6615	20	331	18.2

* N.K. *
~~28/3/~~57

Crude sum of squares for dams, repeating each dam's record

$$= (176)^2 + (222)^2 + (222)^2 + (82)^2 + (32)^2 + (66)^2 + (66)^2 + (20)^2 = 186405 \text{ ....(Col. 5)}$$

Correction term for sum of squares for dams -

$$= \frac{(\text{Total for dams})^2}{\text{Number of pairs}} = \frac{(1067)^2}{8}$$

$$= 142311 \text{ --------------------- (Col. 6)}$$

Corrected sum of squares for dams

$$= \text{Crude S.S. - Correction term}$$

$$= 186405 - 142311$$

$$= 44094 \text{ -------------------------(Col. 7)}$$

The corrected sum of squares for daughters is obtained in the same manner.

Crude sum of squares for daughters

$$= (134)^2 + (118)^2 + (174)^2 + (136)^2 + (66)^2 + (110)^2 + (131)^2 + (194)^2 = 151905 \text{ .... (Col. 8)}$$

Correction term

~~xxk41246xxx~~

$$= \frac{(1063)^2}{8}$$

$$= 141246 \text{ .... (Col. 9)}$$

Corrected sum of squares for daughters

$$= 151905 - 141246$$

$$= 10659 \text{ .... (Col.10)}$$

The corrected sum of product is calculated next using the product of the yields for each pair.

Crude sum of products

$$= (176 \times 134) + (222 \times 118) + (222 \times 174) + (82 \times 136) + (32 \times 66) + (66 \times 110) + (66 \times 131) + (201 \times 194)$$

$$= 156572 \text{ ....(Col.11)}$$

Correction term for the sum of products

$$\frac{(\text{Total for dam})\ (\text{Total for daughter})}{\text{Number of pairs}}$$

$$= \frac{(1067) \times (1063)}{8}$$

= 141778 ....(Col. 12)

Corrected sum of products

= 156572 - 141778

= 14794 ....(Col. 13)

If the records are available for only one sire and the number of daughter-dam pairs is large, say 50 or more, the intra-sire regression may be estimated as the quotient obtained by dividing the corrected sum of products by the corrected sum of squares for dams. But such a large number of pairs from a single sire is hardly likely to be available, the common situation being a number of sires from the same herd with much fewer daughter-dam pairs each. In such cases the corrected sums of products and the sums of squares for dams may be pooled over all sires. The intra-sire regression coefficient may then be obtained as ratio of the pooled sums of products to that for the pooled sums of squares for the dams. For the Hariana herd taken for illustration, here, data in respect of ten sires were available. Computations made for these sires in the manner explained above are presented in Table II. The estimate of the intra-sire regression coefficient (b) obtained from the pooled data is

$$= \frac{94065}{320726}$$

= 0.2933

The corrected daughter average index for Sire I is now obtained as below:-

Average for dams $= \frac{1067}{8}$

= 133.4 .... (Col. 14)

Deviation of the average for dams from the herd average

= 133.4 - 130

= + 3.4 ....(Col. 15)

Correction for the effect of dams

= (Dams av. - herd av.) X regression coefficient

= + 3.4 X 0.2933

= + 1.0 .... (Col. 16)

Average for daughters = $\frac{1063}{8}$

= 132.9 .... (Col. 17)

Corrected Index

= Daughter av. - correction

= 132.9 - 1.0

= 131.9 .... (Col.18)

The following are the steps for computing the standard error of the index.

Residual sum of squares

= (Corrected S.S. for daughters) + (corrected S.S. for dams) X $(b)^2$ - (Corrected S.P) X 2b

where S.S. and S.P denote the sum of squares and products respectively.

= 10659 + (44094) X $(0.2933)^2$ - 14794 X 2 X 0.2933

= 5774 .... (Col. 21)

The deviser for the residual sum of squares

= n X (n - 1)

= 8 X 7

= 56. . . . . (Col. 22).

Variance of the Sire Index

= $\frac{\text{Residual sum of squares}}{\text{Diviser in Col.(22)}}$

= $\frac{5774}{56}$

= 103 .... (Col. 23)

Standard error of the index

= $\sqrt{\text{Variance}}$

= = $\sqrt{103}$ = 10.1 ...... (Col.24)

The method given for the calculation of the standard errors of the indices is a simplified approximation to the exact procedure, as the component term in the sampling variance due to the sampling nature of the regression estimate is neglected. The extent of under-estimation in the standard errors of the indices will, however, decrease with an increase in the volume of data on which the regression coefficient is based. Experience in the analysis of breeding data at the Indian Council of Agricultural Research suggests that the bias involved in using the simplified method is negligible, being of an order less than one per cent, unless the data available for estimating the intra-sire regression is very scanty. The component of the sampling variance of the regression coefficient is

$$\frac{\text{Residual S.S.}}{n-1} \times \frac{(\text{Dam's Av.} - \text{Herds' av'})^2}{\text{Corrected S.S. for dams from total line}}$$

and should be added to column (23) in order to obtain the exact variance of the regression coefficient. For example, the correction needed for the estimate of variance, 103, corresponding to Sire I in column (23), Table II is

$$\frac{5774}{7} \times \frac{(+3.4)^2}{320727}$$ which is less than o.5