SELECTION INDEX

Selection index is a method of multitrait selection in which an individual is selected on the basis of an index value calculated by adding the scores for each trait according to their relative economic value, heritability, phenotypic and genetic correlation with other traits using the multiple regression equation. This index value is an index of the aggregate breeding value or net merit or genetic worth of an individual for many traits. It is also known as total score method. The individuals with the highest scores are kept for breeding purposes.

The first application of the selection index to plant breeding was by Smith (1936) using the discriminant function of Fisher (1936), and first to animals by Hazel (1943). Selection index is also known as discriminant function, since it is used in discriminating individuals with high and low scores.

Selection index uses the principle of multiple regression equation to predict a dependent variable from two or more independent variables.

Let P_i is the phenotypic value of character i for an individual

G_i = Genotypic value of character i

a_i = Relative economic weight given to i^{th} character

H = Aggregate genetic worth or net merit of an individual

$$H = \sum_{i=1}^{n} a_i G_i = a_1 G_1 + a_2 G_2 + a_3 G_3 + \text{---------} + a_n G_n$$

Since G_i's are unknown, H cannot be used as a criterion of selection. Under the circumstances, selection has to be based on some function of the observed values of the various characters which are linearly related to the average genetic worth of individual.

Obviously, it is simplest to have a linear function, I given by

$$I = \sum_{i=1}^{n} b_i P_i = b_1 P_1 + b_2 P_2 + b_3 P_3 + \text{---------} + b_n P_n$$

I = Index value

b_i = partial regression coefficient of H on I for character i

P_i = Phenotypic value of ith character

By finding the value of b_i's such that the function 'I' may best discriminate those individuals who have the greatest genotypic economic score, H. in other words, selection based on 'I' should be as nearly good as based on H, had it been known.

The 'I' will reflect the H accurately when the correlation between these (r_{HI}) would be maximum.

When r_{HI} is maximum, then

$H = I$

$$\sum a_i G_i = \sum b_i P_i$$

In Matrix notations

$$aG = bP$$

$$b = P^{-1}aG$$

b = column vector of partial regression coefficient to be calculated
P^{-1} = Inverse of phenotypic variance covariance matrix
a = column vector of economic weights
G = genotypic variance covariance matrix

In form of equations, the number of equations will be equal to number of traits under selection. For 2 traits,

$$b_1P_{11}+b_2P_{12} = a_1G_{11}+a_2G_{12}$$
$$b_1P_{12}+b_2P_{22} = a_1G_{12}+a_2G_{22}$$

In matrix format:

$$\begin{pmatrix} b_1 \\ \\ b_2 \end{pmatrix} = \begin{pmatrix} P_{11} & P_{12} \\ \\ P_{21} & P_{22} \end{pmatrix}^{-1} \begin{pmatrix} a_1 \\ \\ a_2 \end{pmatrix} \begin{pmatrix} G_{11} & G_{12} \\ \\ G_{21} & G_{22} \end{pmatrix}$$

After calculating the value of b's, it is fitted in the mathematical description of function (I)

$$I = b_1P_1 + b_2P_2+b_3P_3+\text{---------}+b_nP_n$$

This fitted equation is called as selection index or total score of individual. The animals are arranged in order of merit on the basis of total score for selection.

Requirements to construct selection index
1. Relative economic value of each trait under selection
2. Genetic and phenotypic variance of all traits
3. Phenotypic and genetic covariance among all traits

Selection index is more efficient method of selection than other methods as

1. It results in more genetic improvement for the time and effort expended in its use.
2. It allows excellence in some traits to make up for slight deficiencies in other while deciding the fate of the individual.

SIRE INDICES

The results of progeny tests are expressed in the form of an index which is the index of the genetic worth of the sire and such an index is known as sire index. It is a mathematical expression of breeding worth of a sire. Based on sire index a numerical value is obtained which indicates the production ability of the sire. Since a number of sires are progeny tested, it requires ranking all sires for their genetic worth so as to select the best sire.

Methods of indexing sires

Different methods for indexing the sire have developed time to time as per the need. A few of them will be discussed here.

Notations used in sire indices:-

$\overline{D}$ = Average of all daughters of a sire under test

C_D = Average of contemporary daughters of the sire

$\overline{M}$ = Average of mates of the sire or dam's average

C_M = Average of contemporary dams of daughters

$\overline{H}$ = Herd average

b = Intra sire regression coefficient of daughters record on dam = $\frac{1}{2} h^2$

Q = Correction factor = $4n / 4 + (n-1) h^2$

n = Number of progeny of the sire under test

B = Breed average

1. Simple Daughter Average Index
Edwards (1932) proposed this sire index as:

$$I_1 = \overline{D}$$

This index is the simplest measure in a single herd under same environment. It provides a sound basis of evaluation but it does not take into account the production level of the mates of the bull (dams of daughters) and hence it is subjected to the bias when production level of mates allotted to different sires are different.

2. Equiparent or Intermediate Index
Hansson (1913) proposed this index and it is also known as Yapp's index (Yapp, 1925) or Mount Hope index because it was first used at Mount Hope Farm in 1928.
This index is:

$$I_2 = 2\overline{D} - \overline{M}$$

This index is based on principle that that two parents contribute equally to the genetic make up of the progeny which is equivalent to saying that the progeny should average midway between the two parents. And hence the name equiparent or intermediate index is given to it.

This index makes adjustment for the variation in production level of the dam (mates). However, it overcorrects the production level of mates allotted to different sires. If inferior dams are allotted to a sire, there is over estimation of sire's breeding value, but in case superior dams than average are allotted, it then underestimate the sire's breeding value. Moreover, it is subject to a relatively high standard error which makes it less useful for comparing the two sires not widely differing in breeding worth.

1. Rice Index

It was given by Rice. It is based on the fact that overall regression of the daughter record on those of their dams is approximately 0.5.

$$I_3 = \frac{(2\bar{D} - \bar{M}) + B}{2}$$

This index is less variable than equiparent index and has same accuracy. It has a breed reference as a reference point.

2. Tomar Index

It was given by Tomar in 1965. It depends on dam daughter comparison with simultaneous use of their contemporary herd averages.

$$I_4 = \bar{D} + (D_e - M_e)$$

$$D_e = \sqrt{\bar{D}} \times C_D$$

$$M_e = \sqrt{\bar{M}} \times C_M$$

3. Corrected Daughter Average Index

This index was developed by Krishnan (1956). It corrects the daughter average for the influence of differential production level of dams sired by bulls on the basis of regression of daughters' record on dams' and it is four times as efficient as the intermediate index.

$$I_5 = \bar{D} - b\,(\bar{M} - \bar{H})$$

The term $b\,(\bar{M} - \bar{H})$ appearing in the index is the correction for the genetic superiority or inferiority of the set of dams allotted to the sire over the herd average.

4. Contemporary Daughter Average Index

This index was proposed by Jain & Malhotra (1970). It takes into account the variation in the number of daughters in the progeny groups and also corrects for period variation in environmental conditions to which the record might have been subjected.

$$I_6 = \overline{H} + bQ\ (\overline{D} - C_D)$$

5. Corrected Contemporary Daughter Average Index

This index was also proposed by Jain & Malhotra (1971). It is an extension of Index I6, and, in addition to adjusting the production records for the differences in the progeny number and period to period variation, adjusts for the unequal production levels of dams mated to different sires.

$$I_7 = \overline{H} + bQ\ [(\overline{D} - C_D) - b\ (\overline{M} - C_M)]$$

6. Dairy Search Index

It was proposed by Sunderasan *et.al* in 1965.

$$I_8 = \overline{H} + \frac{n}{n+12}\ [(\overline{D} - C_D) - b\ (\overline{M} - C_M)]$$

7. Least square constants

The sire constants are obtained by least square techniques which adjust the data for all the environmental effect including the non orthogonality in data. These sire constants are used in getting the sire index as:

$$I_9 = \overline{H} + \frac{2nh^2}{4 + (n-1)\ h^2}\ (S_i)$$

8. Maximum likelihood method and REML

The least square method minimizes the error variance whereas the maximum likelihood method estimates the parameters by maximizing the logarithm of the likelihood function. The likelihood function is the likelihood of simultaneous occurrence of observations and is generally the product of density function of the observations. However, the ML estimates are biased because no account is taken of the degree of freedom in estimating the variance components.

The ML method was improved by a method known as restricted maximum likelihood (REML) which takes care of the bias in estimates as well as avoids negative estimates of component of variance. The variance components by REML are estimated based on residuals calculated after fitting by ordinary least squares from the fixed effect part of the model. This maximizes a marginal likelihood function. This is also called as residual maximum likelihood or marginal maximum likelihood.

9. Best Linear Unbiased Prediction (BLUP) Method

The BLUP method of estimation was developed by Henderson (1949, 1973). This method is more powerful than the conventional selection index approach. This provides directly comparable estimates of the average breeding value of groups of animals born in different years. It takes onto account the complications of non random mating, sire from more than one herd, environmental trends over time, herd differences for breeding value of dams and bias due to selection. Thus BLUP takes account of the fixed effects such as herd-year-0season and it is applicable when mixed models are used. The BLUP eliminates the non genetic biases in estimating B.V. not only this has this but it also remove the genetic biases taking into account the effects of non random mating, genetic merit of dams and the selection. It is the best method of sire evaluation.

10. Other Methods For Multi Herds

Some methods for evaluation of sires used in more than one herd have also been proposed. These are Stable Mate Daughter Average Index, Contemporary Daughter Average Index, Corrected Stable Mate Daughter Average Index and Simplified Regressed Least Squares Method (SRLS).

****