Theor. Appl. Genet. 52, 91-94 (1978)

Some Statistical Properties of an Index of Multiple Traits*

A.W. Nordskog
Department of Animal Science, Iowa State University, Ames, Iowa (USA)

Summary. Hazel (1943) defined a selection index that maximizes the correlation, R_{IH}, between the index, $I = \Sigma b_i x_i$ and its aggregate genetic value, $H = \Sigma a_i g_i$, where the b_i's are the derived coefficients of the observed traits, x_i, the a_i's are their relative economic values and the g_i's are their respective breeding values. This is called an optimum index. The expected value of the index is not H but rather $K = \Sigma b_i g_i$. The ratio of R_{HI}/R_{KI} is always less than or equal to 1.0. With selection on I, the ratio of the change in K to the change in H is the true heritability of the indez, that is, $h_I^2 = \Delta(K)/\Delta H$. This is not the same as R^2_{IH}, which serves only as the predictor of the change in H with selection on I. If the index, I, itself is considered as a unit trait, studies can then be made of correlated response in I when selection is based only on a single trait in the index. I is then called a performance index. This approach provides for additional insight into the question of the failure of selection to make gains in total performance over many generations.

Key words: Selection index – Performance Index – I

Introduction

An index of multiple traits may be defined either as a selection index or as a performance index. A selection index is designed to maximize genetic-economic merit or aggregate breeding value (Hazel 1943) for multiple-traits among individuals in a population. A performance index is a single-valued measure of the multiple-trait performance of a population. Although the same identical values on individuals in a population may be used both as a selection index and as a performance index, their concepts differ, and therefore the statistical properties of each will differ. In particular, each has its own distinct breeding value so that for a given amount of selection on the index, the expected changes in the two breeding values will differ. Furthermore, the classic definition of heritability does not hold for the selection index although it is valid for the performance index.

Definitions and Basic Concepts

Before proceeding with proof of these assertions, it will be useful to set down certain basic definitions and assumptions.

A unit trait 'x' is composed of two statistically independent parts, an additive component, g, and an environmental component, e, so that x = g + e. With the observations on x measured as deviations from the population mean, the expected means, variances, and covariances are

$$\epsilon(x) = \epsilon(g) = \epsilon(e) = 0$$

$$VAR(x) = V_P(x) = V_G(x) + V_E(x)$$

$$Cov_G(x) = Cov_{G\,P}(x) = V_G(x)$$

where the subscripts G, E, and P denote genotype, environment, and phenotype, respectively.

Heritability is defined as a variance ratio or equivalently in terms of regression or correlation;

$$\text{Regression: } \beta_{G\,P}(x) = \frac{Cov_{G\,P}(x)}{V_P(x)} = \frac{V_G(x)}{V_P(x)} = h_x^2$$

$$\text{Correlation: } R^2_{G\,P}(x) = \frac{Cov^2_{G\,P}(x)}{[V_G(x)V_P(x)]} = \frac{V_G(x)}{V_P(x)} = h_x^2$$

Thus, the squared correlation between breeding value and phenotype is equivalent to the regression of breeding value on phenotype, which, in turn, is equivalent to the ratio of genetic to phenotypic variance for breeding value.

---

* Journal Paper No. J-8176 of the Iowa Agriculture and Home Economics Experiment Station, Project 1711.

0040-5752/78/0052/0091/$ 1.00

For multiple traits $x_1, x_2 \dots x_n$, the covariance are,

$$V_G(x_i) = g_{ii}, V_P(x_i) = P_{ii}$$

$$Cov_G(x_i x_j) = g_{ij}, Cov_P(x_i x_j) = P_{ij}$$

Turning now to a selection index, I, and its economic genetic breeding value, H, which I is designed to estimate:

$$I = b_1 x_1 + - - b_i x_i + - - b_n x_n = \sum_{i=1}^{n} b_i x_i$$

$$H = a_1 g_1 + - - a_i g_i - - + a_n g_n = \sum_{i=1}^{n} a_i g_i$$

The x_i are the unit traits in the index, each weighted by a coefficient, b_i. The g_i are the corresponding unit-trait breeding values, each weighted by its relative economic value, a_i.

In matrix terms, the genetic variance of the index is,

$$V_G(I) = V_G(\Sigma b_i x_i) = b'Gb$$

where $b' = (b_1\ b_2, \dots, b_n)$ is the vector of index coefficients, and $G = (g_{ij})$ is the matrix of genetic variances and covariances. Similarly, $V(H) = a'Ga$. The phenotypic variance of the index is $V_P(I) = V_P(\Sigma b_i x_i) = b'Pb$, where P is the phenotypic variance-covariance matrix.

To derive the optimum selection index, the correlation between I and H is maximized, leading to the equations $Pb = Ga$. For the solution, $b = P^{-1} Ga$ so that

$$V(I) = b'Pb' = b'Ga = Cov(H,I)$$

$$\text{and } R^2_{HI} = \frac{Cov^2(H,I)}{V(H)V(I)} = \frac{V(I)}{V(H)} = \frac{b'Pb}{a'Ga} = \beta_{HI}\ \beta_{IH}$$

This quantity may be thought of as the heritability of the index in the sense that response to selection on I predicts genetic change in H. In fact, $\beta_{HI} = \frac{COV(H,I)}{V(I)} = 1$ so that $\beta_{IH} = R^2_{HI}$ as already shown.

Now consider the same index, I, as a unit trait or performance index with breeding value, K. The quantity K is a function of the breeding values of the x_i weighted by the b_i coefficients and not by the a_i coefficients, as in the case of the selection index, so that

$$K = b_1 g_1 + \dots + b_n g_n = \sum_{i=1}^{n} b_i g_i,$$

$$V(K) = Cov(K,I) = b'Gb = V_G(I)$$

and

$$h^2_{(I)} = \frac{V(K)}{V(I)} = \frac{V_G(I)}{V_P(I)} = \frac{b'Gb}{b'Pb}$$

$$\text{Also, } \beta_{KI} = \frac{Cov(K,I)}{V(I)} = \frac{V_G(I)}{V_P(I)} = h^2 = R^2_{KI}$$

Thus, the heritability of a performance index agrees with the basic definitions and assumptions previously stated because it is a unit trait. This is not so for the selection index. Note, in particular, that $R^2_{KI} = h^2_I$ is the true heritability of the index, but this is not the same as R^2_{HI}, the predictor of genetic change in H when selection is on I.

In examining certain selection indexes derived from several chicken populations, R^2_{HI} was always found to be less than R^2_{KI}. The ratio of R^2_{HI}/R^2_{KI} in matrix form is,

$$\frac{V(K)V(I)}{Cov^2(K,I)} \cdot \frac{Cov^2(H,I)}{V(H)V(I)} = \frac{(b'Ga)^2}{(b'Gb)(a'Ga)}$$

By application of the Cauchy-Schwartz inequality, it can be shown that $R_{HI}/R_{KI} \leqslant 1$. Thus, the correlation of the index with H is always less than or equal to the correlation of the index with K.

The response in H to a given amount of selection on I, $i\sigma_I$, is

$$\Delta H = i\,\beta_{HI}\sigma_I = i\ Cov(H,I)/\sigma_I = i(b'Ga)/(b'Pb)^{1/2} \qquad (3)$$

and the concomitant response in K is

$$\Delta K = i\,\beta_{KI}\sigma_I = i\ Cov_{(K,I)}/\sigma_I = i(b'Gb)/(b'Pb)^{1/2} \qquad (4)$$

The ratio of response in (4) to the response in (3) is equal to

$$\frac{\Delta(K)}{\Delta(H)} = \frac{\beta_{KI}}{\beta_{HI}} = \frac{b'Gb}{b'Ga} = \frac{b'Gb}{b'Pb} = h^2_I$$

With selection on I giving a unit of change in economic breeding value, H, the change in breeding value, K, is h^2_I.

Finally, one might pose the question: is the index I, which is optimal to improve H, also optimal to improve K?

Economic merit for H is $\Sigma a_i g_i$ and the best index is $I(H) = b'x$ in terms of vectors b and x. By analogy, the measure of merit for K is $\Sigma b_i g_i$ and the best index is $I(K) = \tilde{b}'x$. This is obtained from the solution of the usual equations, $P\tilde{b} = Gb$, except that here the economic values (a's) are replaced by b's. The b is $P^{-1} Ga$ and, therefore, $P\tilde{b} = GP^{-1} Ga$ and $\tilde{b} = P^{-1} GP^{-1} Ga$. Clearly the vector b is not the same as the vector $\tilde{b}$ and, therefore, $b'x \neq \tilde{b}'x$. Hence, the index $b'x$ is not optimal for the improvement of K.

Example

To gain some insight into the question of performance plateaus when selection involves multiple traits, a breeder might look at genetic changes, not only in the component traits of performance, but also in the total score (index) of the performance traits.

Selection indexes were constructed from parameter estimates (Nordskog 1974) on 4 traits (body weight, egg weight, part-record and full-record egg production) in 5 Leghorn lines. The variances, b'Pb, a'Ga, and b'Gb, were then computed, from which $R^2_{HI} = b'Pb/a'Ga$ and $h^2_I =$

Table 1. Variances, Correlations and Heritabilities of Selection (or Performance) Indexes for 4 Traits in each of 5 Populations

Line	Selection Criterion	V(I)= b'Pb	V(H)= a'Ga	V(K)= b'Gb	R^2_{HI}	h^2_I	R^2_{HI}/h^2_I
Q	Random	38.41	63.97	31.80	.60	0.83	0.72
A	High Egg Prod.	25.23	65.72	12.38	.40	0.49	0.82
C	Low Body Wt.	17.09	51.12	7.09	.33	0.41	0.80
D	High Egg Wt.	9.11	37.34	2.51	.24	0.28	0.86
G	L-BW, H-EW	27.29	63.95	14.51	.43	0.53	0.81
	Av. or ratio	23.43	55.82	13.66	.42	0.58	0.72

b'Gb/b'Pb were derived. These are given in Table 1.

Note that R^2_{HI} is less than h^2_I in each case. The ratios of these range from 0.72 to 0.86. The quantity, b'Gb, represents the aggregate genetic variance of I when thought of as a performance index. Values of V(K) = b'Gb range from 2.51 to 31.80. The highest was for Line Q selected at random. The lowest was for Line D selected for large egg size. The evidence indicates that selection, even for a single trait, has reduced genetic variance of the total score or index. Differences in V(K) are reflected not only by differences in h^2_I, but also by differences in R^2_{HI}, as one would expect because these are strongly correlated.

Taking an average from the 5 populations sampled, $R^2_{HI} = 0.42$ implies that 42% of the variance in economic merit (H) is predictable from the index. On the other hand, $R^2_{KI} = h^2_I = 0.58$ implies that 58% of the variance of the index itself, when considered as a performance index, is of genetic origin. This value is in accord with the classic definition of heritability.

Discussion

The index, I, can serve two purposes. First and foremost, I is a predictor of H. This is the original purpose for which I was invented. Secondly, I can be used as a performance index. This permits, for example, examining the question of genetic change in net or total performance resulting from any particular selection program based on one or more traits. Thus, the net change in I can be viewed as a correlated response. Because selection for a single trait over successive generations may alter the genetic parameters, total performance may be more resistant to change than one or more single traits making up the index of total performance. For example, effective selection for high rate of egg production is known to reduce body size and egg size as correlated responses (Kinney et al. 1970; Nordskog et al. 1974). As a result, an index of total performance involving these three traits seems not to have changed very much (Nordskog 1974).

The quantity, V(I)/V(H) although valid and useful for prediction of gains in economic merit (H) from selection on I, has certain unique but peculiar properties. For example, if H is defined as the 'aggregate genotype' which Hazel (1943) defined it as, then by analogy should not I be defined as the 'aggregate phenotype'? If so, does V(I)/V(H) = $R^2_{HI} = \beta_{IH}$ imply that phenotype is being predicted from genotype? Obviously this doesn't make sense because we wish to predict H from I. Because $R^2_{HI} = \beta_{HI}\beta_{IH} = \beta_{IH}$, it is clear that $\beta_{HI} = 1$. The index is so designed such that one unit of change in I is expected to change H by one unit.

The index I is a unique prediction quantity. When used as a selection index it is a phenotypic estimator of the breeding value, (H), of multiple traits. On the other hand, because $V_I \leqslant V_H$, it could lead the novice to the false conclusion that 'aggregate phenotypic variance' is less than or equal to 'aggregate genetic variance'. Hazel's (1943) use of the term 'aggregate genotype' to define H seems not to have been the best choice.

I is a predictor of H, a hypothetical genetic quantity. However, I can be regarded as a phenotypic observation on an individual only when it is thought of as a unit trait, i.e., a performance index. Hence, we require the definition of $K = b_1g_1 + b_2g_2$, etc. so that V_K = b'Gb is the valid genetic variance of the index and therefore V(I) = b'Pb is now clearly the phenotypic variance.

The quantities V(H) = a'Ga and V(K) = b'Gb = $a'GP^{-1}GP^{-1}Ga$ are obviously closely correlated because each contains common elements. Therefore, for all practical purposes, if one directs hus inquiry to the genetic variance of the index, either V(H) or V(K) could be used even though V(K) is theoretically the correct one.

Acknowledgment

I greatly appreciate the helpful discussions I have had with my colleagues at ISU, Drs. O. Kempthorne and Dave Cox. I also received a helpful note on interpretation from Dr. W.G. Hill, Institute of Animal Genetics, Edinburgh.

Literature

Hazel, L.N.: The genetic basis for constructing selection indexes. Genetics **28**, 476-490 (1943)

Kinney, T.B.; Bohren, B.B.; Craig, J.V.; Lowe, P.C.: Responses to individual, family or index selection for short term rate of egg production in chickens. Poultry Sci. **49**, 1052-1064 (1970)

Nordskog, A.W.: Multi-trait selection and egg production in chickens. Proc. 1st World Cong. Genet. Appl. to L.S. Prod. **1**, 917 (1974)

Nordskog, A.W.; Tolman, H.S.; Casey, D.W.; Lin, C.Y.: Selection in small populations of chickens. Poultry Sci. **53**, 1188-1219 (1974)

Received September 12, 1977
Accepted November 28, 1977
Communicated by H. Abplanalp

Dr. A.W. Nordskog
Department of Animal Science
Iowa State University
Ames, Iowa 50011
(USA)

Editor's Note: This paper and one closely similar (Lin and Allaire T.A.G. **51**, 1-3 (1977) were independently derived and reviewed by different corresponding editors. The earlier paper, however, lacks numerical examples and a possible application as discussed in the present paper.