A method to construct confidence interval for expected response to multi-trait selection

G. C. C. Tai

Agriculture Canada Research Station, P.O. Box 20280, Fredericton, New Brunswick E3B 4Z7, Canada

Summary. Confidence intervals are constructed for the expected responses to three types of multi-trait selection. The influence of numbers of replicates and genotypes used in a progeny test experiment on the precision of response of multi-trait selection is discussed based on the structure of the established intervals. Special attention is paid to the characteristics of the intervals constructed for the conventional least square selection indices.

Key words: Multi-trait selection – Confidence interval – Index selection

Introduction

One of the most serious problems in using index selection (Smith 1936; Hazel 1943) in breeding work is that the true population parameters are usually unknown and have to be estimated from samples.

The study of effect of sampling errors on the efficiency of index selection is difficult and often involves complicated statistical procedures (see e.g., Williams 1962a; Harris 1964; Hayes and Hill 1980). Tallis (1960) obtained a large sample variance of the index weights and predicted genetic gains. Harris (1963, 1964) examined the concept of progress from index selection and studied the joint distribution of the expected progress from selection for a calculated index and the predicted progress from this index. Hayes and Hill (1980) elaborated further on the definition of response to selection and study the sampling properties of the index weights by a reparameterization of the phenotypic and genetic covariances used in constructing the selection indices.

The aim of the present report is to propose a method to construct a confidence interval for the expected response to three types of multi-trait selection. The interval is then investigated for its various implications in selection work.

General theory

Consider a set of n variables v_i, $i = 1, 2, \ldots n$, which follows multivariate normal distribution $MN(\mu, \Sigma)$ with

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \cdot \\ \cdot \\ \cdot \\ \mu_n \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{bmatrix}$$

in which μ_i is the mean, $\sigma_{ii} (= \sigma_i^2)$ the variance of i-th variable, and σ_{ij} the covariance of i-th and j-th variables. Let $\mathbf{d}$ be a $(n \times 1)$ vector of constants, i.e., $\mathbf{d}' = [d_1 \, d_2 \ldots dn]$. A new composite variable $z = \mathbf{d}' v$ should have mean $\mu_z = \mathbf{d}' \mu$ and variance $\sigma_z^2 = \mathbf{d}' \Sigma \, \mathbf{d}$. Also, for a random sample of size m,

$$\frac{\mathbf{d}' \mathbf{S} \mathbf{d}}{\mathbf{d}' \Sigma \mathbf{d}} = \chi_{df}^2, \quad df = m - 1$$

where $\mathbf{S}$ is the matrix of sample sum of squares and products. A $(1 - \alpha)$ confidence interval for $\mathbf{d}' \Sigma \mathbf{d}$ can be obtained as

$$P \left\{ \frac{\mathbf{d}' \mathbf{S} \mathbf{d}}{\chi_{\alpha/2:\,df}^2} \leq \mathbf{d}' \Sigma \mathbf{d} \leq \frac{\mathbf{d}' \mathbf{S} \mathbf{d}}{\chi_{1-\alpha/2:\,df}^2} \right\} \geq 1 - \alpha \tag{1}$$

where $\chi_{\alpha/2:\,df}^2$ and $\chi_{1-\alpha/2:\,df}^2$ are tabulated χ^2 value with $m - 1$ degrees of freedom at the probability level $\alpha/2$ and $(1 - \alpha/2)$, respectively.

Now assume that data of n traits are obtained from each of m genotypes tested in an experiment replicated r times according to a randomized complete block design. Let x_{ijk} be the observed value of j-th trait of

i-th genotype in k-th block,

$$x_{ijk} = u_j + b_{jk} + g_{ij} + \varepsilon_{ijk}$$

$$i = 1, \ldots, m; \ j = 1, \ldots, n; \ k = 1, \ldots, r,$$

where u, b, g and ε are, respectively, the population mean, and effects due to k-th block, i-th genotype and i k-th error deviate for the j-th trait. The phenotypic effect of j-th trait of i-th genotype is then

$$P_{ij} = \bar{x}_{ij.} - \bar{x}_{.j.} = g_{ij} + e_{ij}$$

where $e_{ij} = \dfrac{1}{r} \sum\limits_{k=1}^{r} \varepsilon_{ijk}$. Let T be the matrix of mean squares and products between genotypes and ε the corresponding matrix of error mean squares and products in the analysis of variance and covariance (VAR-COV) table for the n traits. Also let P, G and E be VAR-COV matrices for the phenotypic, genotypic and error effects,

$$P = \begin{bmatrix} \sigma_{p11} & \sigma_{p12} & \cdots & \sigma_{p1n} \\ \sigma_{p21} & \sigma_{p22} & \cdots & \sigma_{p2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \sigma_{pn1} & \sigma_{pn2} & & \sigma_{pnn} \end{bmatrix}$$

$$G = \begin{bmatrix} \sigma_{g11} & \sigma_{g12} & \cdots & \sigma_{g1n} \\ \sigma_{g21} & \sigma_{g22} & \cdots & \sigma_{g2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \sigma_{gn1} & \sigma_{gn2} & & \sigma_{gnn} \end{bmatrix}$$

$$E = \begin{bmatrix} \sigma_{e11} & \sigma_{e12} & \cdots & \sigma_{e1n} \\ \sigma_{e21} & \sigma_{e22} & \cdots & \sigma_{e2n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \sigma_{en} & \sigma_{en2} & & \sigma_{enn} \end{bmatrix}$$

P, G and E are estimated by

$$\hat{P} = \frac{1}{r} T, \quad \hat{G} = \frac{1}{r} (T - \varepsilon) \quad \text{and} \quad \hat{E} = \frac{1}{r} \varepsilon .$$

Three types of expected response to selection are considered in the present paper:

1 Achieved response

Let a and b be two $(n \times 1)$ vectors of constants whose elements are two sets of weights associated with the n traits. a is used to weigh the genotypic effects of n traits of a genotype to obtain an aggregate genotypic value, i.e., $h_i = \sum\limits_{j=1}^{n} a_j g_{ij}$, whereas b to obtain an index $I_i = \sum\limits_{j=1}^{n} b_j p_{ij}$. The achieved response to selection based

on I_i is measured by $R_a = i \, a' \, G \, b \, (b' \, P \, b)^{-1/2}$ where i is the intensity of selection.

2 Correlated response

The response of a trait to selection for another trait in the population is called correlated response (Falconer 1982). Let $C R_{kl}$ be the correlated response of k-th trait to selection for l-th trait. We have $C R_{kl} = i \, \sigma_{gkl} (\sigma_{pll})^{-1/2}$. Let a and b be two vectors with null elements but the k-th and l-th elements, respectively, taking the value of unity. Then, $C R_{kl} = R_a$.

3 Maximum attainable response

In conventional least square index selection, a is usually a set of "economic weights" whereas b is obtained by $b = P^{-1} G a$ (Smith 1936). The response to selection based on the index $I_i = \sum\limits_{j=1}^{n} b_j P_{ij}$ is measured by $R = i \, (b' \, P \, b)^{1/2}$. This represents the maximum attainable response when the genotypic effects are weighed by a.

It is noted R_a, $C R$ and R can only be estimated using $\hat{G}$ and $\hat{P}$. Also, R_a and $C R$ can be either positive or negative measures, whereas $R \geq 0$ when the intensity of selection (i) is positive.

Confidence intervals for expected response to multi-trait selection

Achieved response

The confidence interval for R_a can be derived by using the confidence intervals for $(a+b)' \, P (a+b)$, $a' \, P \, a$, $b' \, P \, b$, $(a+b)' \, E (a+b)$, $a' \, E \, a$, and $b' \, E \, b$. Let the six intervals with confidence limit $(1-a)$ be (L_1, U_1), (L_2, U_2), (L_3, U_3), (L_4, U_4), (L_5, U_5) and (L_6, U_6), respectively. The $(1-\alpha)$ interval (L_1, U_1), for example, can be derived by putting $d = (a+b)$, $S = (m-1) \, \hat{P}$, and $df = (m-1)$ in (1), i.e.,

$$P \left\{ \frac{(m-1) \, (a+b)' \, \hat{P}(a+b)}{\chi^2_{\alpha/2: \, m-1}} \leq (a+b)' \, P (a+b) \right.$$

$$\left. \leq \frac{(m-1) \, (a+b)' \, \hat{P}(a+b)}{\chi^2_{1-\alpha/2: \, m-1}} \right\} \geq 1-\alpha .$$

Other five $(1-\alpha)$ intervals can be similarly derived. Since both P and E are positive semi-definite (see Searle 1966), the following two inequalities can be set up:

$$P \{ L_a \leq (a+b)' \, P (a+b) + a' \, E \, a + b' \, E \, b \leq U_a \} \geq 1 - \alpha$$

$$P \{ L_b \leq (a+b)' \, E (a+b) + a' \, P \, a + b' \, P \, b \leq U_b \} \geq 1 - \alpha$$

where $L_a = L_1 + L_5 + L_6$, $U_a = U_1 + U_5 + U_6$, $L_b = L_2 + L_3 + L_4$ and $U_b = U_2 + U_3 + U_4$. Let $A =$

$a' P a + (a+b)' E (a+b)$, $B = b' P b$ and $C = a' G b$ and note that $P = G + E$. We have

$$(a+b)' P (a+b) + a' E a + b' E b = A + B + 2 C$$
$$(a+b)' E (a+b) + a' P a + b' P b = A + B.$$

A Bonferroni inequality can then be formulated as follows (see Graybill 1976):

$$P \{L_a \le A + B + 2C \le U_a, \ L_b \le A + B \le U_b,$$
$$L \le C^2/B \le U\} \ge 1 - 3\alpha.$$

That is,

$$P \{L_a \le (a+b)' P (a+b) + a' E a + b' E b \le U_a,$$
$$L_b \le (a+b)' E (a+b) + a' P a + b' P b \le U_b,$$
$$L \le (a' G b)^2/b' P b \le U\} \ge 1 - 3\alpha. \tag{2}$$

The relationship in (2) can be used to obtain a $(1 - 3\alpha)$ confidence interval for $b' P b$ according to the method given by Williams (1962b):

$$P \{L_c \le b' P b \le U_c\} \ge 1 - 3\alpha. \tag{3}$$

The structures of L_c and U_c are known,

$$L_c = (m - 1) \ b' \hat{P} b/\chi^2_{3\alpha/2: m-1},$$
$$U_c = (m - 1) \ b' \hat{P} b/\chi^2_{1-3\alpha/2: m-1}.$$

Therefore, (2) and (3) can be used to find L and U. Let l_a, l_b and l_c be the observed lower limits of L_a, L_b and L_c, respectively. We have

$$\begin{cases} l_a = (a+b)' P (a+b) + a' E a + b' E b \\ l_b = (a+b)' E (a+b) + a' P a + b' P b \\ l_c = b' P b \\ l = (a' G b)^2/b' P b. \end{cases}$$

Solving the equations for l,

$$l = (l_a - l_b)^2/4 l_c.$$

The general expression of L is then

$$L = \left[\frac{(m-1) \ a' \hat{P} b}{\chi^2_{\alpha/2: m-1}} - \frac{(r-1) \ (m-1) \ a' \hat{E} b}{\chi^2_{\alpha/2: (r-1) (m-1)}} \right]^2$$
$$\cdot \left[\frac{(m-1) \ b' \hat{P} b}{\chi^2_{3\alpha/2: m-1}} \right]^{-1}. \tag{4}$$

Similarly,

$$U = \left[\frac{(m-1) \ a' \hat{P} b}{\chi^2_{1-\alpha/2: m-1}} - \frac{(r-1) \ (m-1) \ a' \hat{E} b}{\chi^2_{1-\alpha/2: (r-1) (m-1)}} \right]^2$$
$$\cdot \left[\frac{(m-1) \ b' \hat{P} b}{\chi^2_{1-3\alpha/2: m-1}} \right]^{-1}.$$

The absolute size of R_a falls between $i L^{1/2}$ and $i U^{1/2}$. Let (L_{Ra}, U_{Ra}) be the $(1 - \alpha)$ confidence interval for R_a. (L_{Ra}, U_{Ra}) can be $(i L^{1/2}, i U^{1/2})$, $(-i L^{1/2}, i U^{1/2})$,

$(-i U^{1/2}, i L^{1/2})$ or $(-i U^{1/2}, -i L^{1/2})$. This can be determined by choosing the shortest one which covers the point estimate R_a. It is interesting to note that $R_a \ge 0$ when $a = b$ and $i > 0$. This is what Williams (1962a) referred to as the base index. In this case $L_R = i L^{1/2}$ and $U_R = i U^{1/2}$. L_R should be set to zero when $(m-1) \ a' P b/ \chi^2_{\alpha/2: m-1} < (r-1) \ (m-1) \ a' E b/\chi^2_{\alpha/2: (r-1) (m-1)}$ whereas both L_R and U_R set to zero when $(m-1) \ a' P b/ \chi^2_{1-\alpha/2: m-1} < (r-1) \ (m-1) \ a' E b/\chi^2_{1-\alpha/2: (r-1) (m-1)}$.

Correlated response

Consider two associated traits 1 and 2. Selection for trait 1 would bring up response of trait 2. The correlated response is measured by $C R_{21} = i \ \sigma_{g12} (\sigma_{p11})^{1/2}$. Let

$$L = \left[\frac{(m-1) \ \hat{\sigma}_{p12}}{\chi^2_{\alpha/2: m-1}} - \frac{(r-1) \ (m-1) \ \hat{\sigma}_{e12}}{\chi^2_{\alpha/2: (r-1) (m-1)}} \right]^2$$
$$\cdot \left[\frac{\hat{\sigma}_{p11}}{\chi^2_{3\alpha/2: m-1}} \right]^{-1},$$
$$U = \left[\frac{(m-1) \ \hat{\sigma}_{p12}}{\chi^2_{1-\alpha/2: m-1}} - \frac{(r-1) \ (m-1) \ \hat{\sigma}_{e12}}{\chi^2_{1-\alpha/2: (r-1) (m-1)}} \right]^2$$
$$\cdot \left[\frac{\hat{\sigma}_{p11}}{\chi^2_{1-3\alpha/2: m-1}} \right]^{-1}. \tag{5}$$

Again (L_{CR}, U_{CR}) can be $(i L^{1/2}, i U^{1/2})$, $(-i L^{1/2}, i U^{1/2})$, $(-i U^{1/2}, i L^{1/2})$ or $(-i U^{1/2}, -i L^{1/2})$. The correct $(1 - \alpha)$ interval should be the shortest one which covers the point estimate $C R_{21}$.

Maximum attainable response

The confidence interval for $b' P b$ can be expressed as

$$P \left[\frac{(m-1) \ b' \hat{P} b}{\chi^2_{\alpha: m-1}} \le b' P b \le \frac{(m-1) \ b' \hat{P} b}{\chi^2_{1-\alpha: m-1}} \right] \ge 1 - 2\alpha$$

when the elements in b are known constants. In the present case, however, b is estimated by $\hat{b} = \hat{P}^{-1} \hat{G} a$. We are thus looking for two estimated vectors b_l and b_u such that $b_l' \hat{P} b_l \le b' \hat{P} b \le b_u' \hat{P} b_u$ and

$$P \left[\frac{(m-1) \ b_l' \hat{P} b_l}{\chi^2_{\alpha: m-1}} \le b' P b \le \frac{(m-1) \ b_u' \hat{P} b_u}{\chi^2_{1-\alpha: m-1}} \right] \ge 1 - 2\alpha. \tag{6}$$

Let I be a $(n \times n)$ unit matrix. We have

$$P b = G a = I G a = \begin{bmatrix} I_1' \ G \ a \\ \cdot \\ \cdot \\ \cdot \\ I_i' \ G \ a \\ \cdot \\ \cdot \\ \cdot \\ I_n' \ G \ a \end{bmatrix}$$

where $\mathbf{I}_i$ is a vector having zero for all its elements but the i-th one unity. Let (L_i, U_i) be the $(1 - 2\alpha)$ confidence interval for $\mathbf{I}_i \mathbf{G} \mathbf{a}$ such that

$$P\left[\sum_{i=1}^{n} b_i L_i \le \mathbf{b}' \mathbf{G} \mathbf{a} \le \sum_{i=1}^{n} b_i U_i\right] \ge 1 - 2\alpha. \qquad (7)$$

L_i and U_i can be derived by establishing the following Bonferroni inequality:

$$P\,[L_{1i} \le (\mathbf{I}_i + \mathbf{a})'\,\mathbf{P}\,(\mathbf{I}_i + \mathbf{a}) + \mathbf{I}_i'\,\mathbf{E}\,\mathbf{I}_i + \mathbf{a}'\,\mathbf{E}\,\mathbf{a} \le U_{1i},$$

$$L_{2i} \le (\mathbf{I}_i + \mathbf{a})'\,\mathbf{E}\,(\mathbf{I}_i + \mathbf{a}) + \mathbf{I}_i'\,\mathbf{P}\,\mathbf{I}_i + \mathbf{a}'\,\mathbf{P}\,\mathbf{a} \le U_{2i}]$$

$$\ge 1 - 2\alpha, \quad i = 1, 2, \ldots, n.$$

Based on the same procedure as described in $(1) \sim (4)$, we obtain

$$L_i = (L_{1i} - L_{2i})/2$$

$$= \left[\frac{(m-1)\,\mathbf{I}_i'\,\hat{\mathbf{P}}\,\mathbf{a}}{\chi^2_{\alpha/2:\,m-1}} - \frac{(r-1)\,(m-1)\,\mathbf{I}_i'\,\hat{\mathbf{E}}\,\mathbf{a}}{\chi^2_{\alpha/2:\,(r-1)\,(m-1)}}\right]$$

$$U_i = (U_{1i} - U_{2i})/2$$

$$= \left[\frac{(m-1)\,\mathbf{I}_i'\,\hat{\mathbf{P}}\,\mathbf{a}}{\chi^2_{1-\alpha/2:\,m-1}} - \frac{(r-1)\,(m-1)\,\mathbf{I}_i'\,\hat{\mathbf{E}}\,\mathbf{a}}{\chi^2_{1-\alpha/2:\,(r-1)\,(m-1)}}\right]$$

$$i = 1, 2, \ldots, n.$$

Let $\mathbf{L}' = [L_1 \ldots L_i \ldots L_n]$ and $\mathbf{U}' = [U_1 \ldots U_i \ldots U_n]$. Then (7) can be rewritten as

$$P\,[\mathbf{b}'\,\mathbf{L} \le \mathbf{b}'\,\mathbf{P}\,\mathbf{b} \le \mathbf{b}'\,\mathbf{U}] \ge 1 - 2\alpha. \qquad (8)$$

Obviously, by examining (8) we can also request that the two vectors $\mathbf{b}_l$ and $\mathbf{b}_u$ fulfil the following relation:

$$P\,[\mathbf{b}_l'\,\mathbf{L} \le \mathbf{b}'\,\mathbf{P}\,\mathbf{b} \le \mathbf{b}_u'\,\mathbf{U}] \ge 1 - 2\alpha. \qquad (9)$$

Comparing (9) with (6), we have

$$\mathbf{b}_l\,\mathbf{L} = \frac{(m-1)\,\mathbf{b}_l'\,\hat{\mathbf{P}}\,\mathbf{b}_l}{\chi^2_{\alpha:\,m-1}}$$

$$\mathbf{b}_u\,\mathbf{U} = \frac{(m-1)\,\mathbf{b}_u'\,\hat{\mathbf{P}}\,\mathbf{b}_u}{\chi^2_{1-\alpha:\,m-1}}$$

$\mathbf{b}_l$ and $\mathbf{b}_u$ are then obtained as

$$\mathbf{b}_l = \chi^2_{\alpha:\,m-1}\,\hat{\mathbf{P}}^{-1}\mathbf{L}/(m-1)$$
$$\mathbf{b}_u = \chi^2_{1-\alpha:\,m-1}\,\hat{\mathbf{P}}^{-1}\mathbf{U}/(m-1) \qquad (10)$$

respectively. Bringing (10) into (9), the confidence interval for $\mathbf{b}'\,\mathbf{P}\,\mathbf{b}$ can be established as:

$$P\left[\frac{\chi^2_{\alpha:\,m-1}\,\mathbf{L}'\,\hat{\mathbf{P}}^{-1}\mathbf{L}}{m-1} \le \mathbf{b}'\,\mathbf{P}\,\mathbf{b} \le \frac{\chi^2_{1-\alpha:\,m-1}\,\mathbf{U}'\,\hat{\mathbf{P}}^{-1}\mathbf{U}}{m-1}\right]$$

$$\ge 1 - 2\alpha. \qquad (11)$$

The confidence interval for the maximum attainable response can be easily constructed using (11). Let (L_R, U_R) be the interval

$$P\,(L_R \le R \le U_R) \ge 1 - 2\alpha$$

where

$$L_R = i\,[\chi^2_{\alpha:\,m-1}\,\mathbf{L}'\,\hat{\mathbf{P}}^{-1}\mathbf{L}/(m-1)]^{1/2}$$
$$U_R = i\,[\chi^2_{1-\alpha:\,m-1}\,\mathbf{U}'\,\hat{\mathbf{P}}^{-1}\mathbf{U}/(m-1)]^{1/2}. \qquad (12)$$

Discussion

An inspection of the intervals in (4), (5) and (11) would reveal the influence of the sizes of r and m on the precision of response to multi-trait selection. Tai (1983) pointed out that the use of a small number of genotypes in an elaborate experiment of many replicates may lead to imprecise genetic information on quantitatively inherited traits. The same conclusion can be made here with regard to response to multi-trait selection. It is important to have a large sample of genotypes tested in an experiment to ensure the precision of the estimate of response to multi-trait selection.

Of the three types of multi-trait selection studied here, the most important one is, of course, the conventional least square index selection. Many theoretical works (Tallis 1960; Williams 1962a; Harris 1964; Sales and Hill 1976a, b; Hayes and Hill 1980) were focused on locating the sampling properties of the index weights and identifying factors which may affect the effectiveness of index selection. The confidence intervals obtained in (12) can obviously be used for the same purpose. The lower confidence limit L_R, for example, can be expressed as follows:

$$L_R = i\,\{k_1\,\mathbf{a}'\,[k_2\,\hat{\mathbf{P}} - k_3\,\hat{\mathbf{E}}]\,\hat{\mathbf{P}}^{-1}\,[k_2\,\hat{\mathbf{P}} - k_3\,\hat{\mathbf{E}}]\,\mathbf{a}\}^{1/2}$$

where

$$k_1 = \chi^2_{\alpha:\,m-1}\,(m-1), \quad k_2 = (m-1)/\chi^2_{\alpha:\,m-1} \quad \text{and}$$
$$k_3 = (r-1)\,(m-1)/\chi^2_{\alpha/2:\,(r-1)\,(m-1)}.$$

This expression is comparable to that of the point estimate R, i.e.,

$$R = i\,\{\mathbf{b}'\,\hat{\mathbf{P}}\,\hat{\mathbf{b}}\}^{1/2} = i\,\{\mathbf{a}'\,(\hat{\mathbf{P}} - \hat{\mathbf{E}})\,\hat{\mathbf{P}}^{-1}(\hat{\mathbf{P}} - \hat{\mathbf{E}})\,\mathbf{a}\}^{1/2}.$$

For $r = 2$, $k_2 = k_3 = (m-1)/\chi^2_{\alpha/2:\,m-1}$ and

$$L_R = i\,\left\{\frac{(m-1)\,\chi^2_{\alpha:\,m-1}}{[\chi^2_{\alpha/2:\,m-1}]^2}\,\mathbf{a}'\,(\hat{\mathbf{P}} - \hat{\mathbf{E}})\,\hat{\mathbf{P}}^{-1}(\hat{\mathbf{P}} - \hat{\mathbf{E}})\,\mathbf{a}\right\}^{1/2}$$

$$= \frac{[(m-1)\,\chi^2_{\alpha:\,m-1}]^{1/2}}{\chi^2_{\alpha/2:\,m-1}} \cdot R.$$

The above relation reveals $L_R \le R$ for $m \ge 2$. The difference between them, however, would decline when increasingly larger m is used. The gap between L_R and R is getting larger when a greater number of replicates is used in an experiment. The trend can be checked by using a larger sample of genotypes in the experiment.

The above expression of L_R also indicates that the relative sizes of economic weights in $\mathbf{a}$ and the esti-

mates of VAR-COV components in **P** and **E** would influence the difference between L_R and R. It is thus possible to use the interval constructed in (12) to study the impact of various factors, i.e., numbers of genotypes and replicates, sizes of estimates of genotypic and phenotypic VAR-COV components, and economic weights, on the precision of index selection.

It is interesting to observe that, for a single trait,

$$L_R = i \left\{ \frac{(m-1)\ \hat{\sigma}_P^2}{\chi^2_{\alpha/2:\ m-1}} - \frac{(r-1)\ (m-1)\ \hat{\sigma}_e^2}{\chi^2_{\alpha/2:\ (r-1)\ (m-1)}} \right\}$$
$$\cdot \left\{ \frac{(m-1)\ \hat{\sigma}_P^2}{\chi^2_{\alpha:\ m-1}} \right\}^{-1/2}$$

according to (12). This is larger than and thus an improved lower limit over the one obtained by Tai (1979, 1983) for single-trait selection, i.e.

$$L_R = \left\{ \frac{(m-1)\ \hat{\sigma}_P^2}{\chi^2_{\alpha:\ m-1}} - \frac{(r-1)\ (m-1)\ \hat{\sigma}_e^2}{\chi^2_{2\alpha:\ (r-1)\ (m-1)}} \right\}$$
$$\cdot \left\{ \frac{(m-1)\ \hat{\sigma}_P^2}{\chi^2_{\alpha:\ m-1}} \right\}^{-1/2} \ .$$

References

Graybill FA (1976) Theory and application of the linear model. Duxbury Press, Belmont, California

Harris DL (1963) Influence of errors of parameter estimation upon index selection. In: Statistical genetics and plant breeding. NAS-NPC Pub 982, pp 491–500

Harris DL (1964) Expected and predicted progress from index selection involving estimates of population parameters. Biometrics 20:46–72

Hayes JF, Hill WG (1980) A reparameterization of a genetic selection index to locate its sampling properties. Biometrics 36:237–248

Hazel LN (1943) The genetic basis for constructing selection indices. Genetics 28:476–490

Sales J, Hill WG (1976a) Effect of sampling errors on efficiency of selection indices. 1. Use of information from relatives for single trait improvement. Anim Prod 22:1–17

Sales J, Hill WG (1976b) Effect of sampling errors on efficiency of selection indices. 2. Use of information on associated trait for improvement of a single important trait. Anim Prod 23:1–14

Searle SR (1971) Linear models. Wiley and Sons, New York

Smith HF (1936) A discriminant function for plant selection. Ann Eugen (London) 7:240–250

Tai GCC (1979) An interval estimation of expected response to selection. Theor Appl Genet 54:273–275

Tai GCC (1983) Confidence intervals of genetic parameters estimated from progeny test experiments in plant breeding. Can J Genet Cytol 25:502–508

Tallis GM (1960) The sampling errors of estimated genetic regression coefficients and the errors of predicted genetic gains. Aust J Stat 2:66–77

Williams JS (1962a) Some statistical properties of a genetic selection index. Biometrika 49:325–337

Williams JS (1962b) A confidence interval for variation components. Biometrika 49:278–281