

Portfolio theory, utility theory and mate selection

S.P. SMITH and K. HAMMOND

Animal Genetics and Breeding Unit, University of New England, Armidale NSW 2351, Australia

Summary

A decision rule is said to be rational if it is consistent with maximising expected utility. Rational decision rules can correctly accommodate objective and subjective criteria. It is shown that portfolio theory, as applied to sire selection problems, is rational. However, the underlying utility function must be linear or exponential and normality assumptions must hold. Explicit use of utility theory and related Bayesian methodology is suggested for other selection problems. A mate selection rule is described which is rational and consistent with the underlying reason for positive assortative mating: to arrange future pedigree information by attempting to create extreme positive individuals. The mate selection rule can be evaluated sequentially using mixed model methodology as this provides for the reuse of information on preassorted relatives.

Key words : Bayesian methods, portfolio theory, positive assortative mating, mate selection, utility theory.

Résumé

Théorie du portfolio, théorie de l'utilité, choix et accouplement des reproducteurs

Une règle de décision est dite rationnelle si elle est cohérente avec la maximisation de l'espérance de l'utilité. Des règles de décision rationnelles peuvent en toute rigueur intégrer des critères objectifs et des critères subjectifs. On montre que la théorie du portfolio, appliquée aux problèmes de sélection et d'accouplement des reproducteurs, est rationnelle. Cependant, il faut que la fonction d'utilité sous-jacente soit linéaire ou exponentielle et que les hypothèses de normalité soient vérifiées. L'utilisation explicite de la théorie de l'utilité et de la méthodologie bayésienne est suggérée pour aborder d'autres problèmes de sélection. On décrit une règle de sélection et d'accouplement des reproducteurs qui est rationnelle et cohérente avec la motivation sous-jacente à l'homogamie positive : organiser l'information généalogique de la génération à venir afin d'y créer des génotypes positifs extrêmes. La règle de sélection et d'accouplement des reproducteurs peut être déterminée de façon continue en utilisant la méthodologie du modèle mixte qui assure la réutilisation de l'information sur des apparentés préalablement assortis.

Mots clés : Méthodes bayésiennes, théorie du portfolio, homogamie positive, sélection et accouplement des reproducteurs, théorie de l'utilité.

I. Introduction

Selection decisions, like most decisions, are affected by subjective beliefs or perceptions. Beliefs are subjective in that they are personal and may change from person to person. Classical quantitative genetics has largely ignored subjectivity. However, the importance of risk in applied genetics has been recognised (e.g., SCHNEEBERGER *et al.*, 1982 ; SMITH, 1985) despite the fact that concerns about risk are mostly subjective (DILLON, 1977).

Risk exists with uncertainty and indeed DILLON (1977) uses the words risk and uncertainty interchangeably. Conditions such as infinite population size or infinite replication tend to remove uncertainty. For example, with infinite populations the response due to truncation selection is deterministic. Alternatively, the average response from an infinite sequence of similar selection schemes is constant. It follows that concerns about risk can be ignored under conditions of infinite population size or infinite replication. Risk can sometimes be ignored under conditions of equal information when equal information implies equal uncertainty across all selection alternatives. In this situation uncertainty is not removed but equalized.

While assumptions like infinite population size, infinite replication and equal information are convenient they need not be realistic. Selection of replacement breeding animals involves decisions by a person or persons on behalf of enterprises. Subjective attitudes cannot generally be ignored. Moreover, because of subjective attitudes about uncertainty, a unique and optimal decision will not usually exist for all decision-makers given the same selection alternatives.

Concern about risk may be mistakenly regarded as purely objective ; due to a belief that risk can be accommodated by :

- 1) extending the planning horizon
- 2) evaluating selection alternatives under condition of infinite replication.

Concern about drift in finite populations is sometimes quantified in this manner (e.g., ROBERTSON, 1960). It may be reasonable to extend the planning horizon, but this extension is liable to make the analysis very difficult. The above view is flawed because infinite replication is a corruption even for long planning horizons.

Portfolio theory has been proposed as an aid to decision-making in dairy sire selection (SCHNEEBERGER *et al.*, 1982). The method is purported to allow for both risk and objective criteria in the decision processes. Given that risk should be considered in decision-making, is portfolio theory an appropriate tool to deal with both objective and subjective criteria? In this paper we consider the issues and show that portfolio theory can lead to rational decisions as defined by the Bayesian. Ramifications will be discussed and we also show how generated theory can be applied to mate selection problems.

II. Utility theory and the axioms of rational behavior

To determine if portfolio theory is indeed rational we need to appeal to a higher realm of meaning ; utility theory and the axioms of rational behavior (BERGER, 1980). Utility theory can be regarded as a component of Bayesian decision theory.

The goal of selection is to increase merit in some future population. This is accomplished by choosing one of several possible selection alternatives (e.g. a group of parents could be one selection alternative). In general, different selection alternatives will induce different outcomes of merit.

Using the terminology of BERGER (1980), we will call a particular outcome of merit a reward and denote it as r . Also, let the set of all possible rewards be denoted by R . When selection decisions are made there is usually uncertainty about which r in R will be realized. Let $P_i(r)$ represent the probability distribution function of possible rewards for the i -th selection alternative. The quantities, $P_1(r)$, $P_2(r)$..., can be regarded as posterior distributions given data, empirical experiences and subjective perceptions. The distributions can be constructed using Bayesian methods (BERGER, 1980). In an animal breeding context, SMITH & ALLAIRE (1985) show how to use mixed model methodology to construct $P_i(r)$ under normality. We will not dwell on the construction of $P_i(r)$ in this paper.

Given $P_1(r)$, $P_2(r)$..., preference among selection alternatives can be quantified in accordance to rationality axioms (SHARPE, 1970 ; BERGER, 1980). The most desirable selection alternative corresponds to the largest of :

$$E_{p_i}[U(r)] = \int_{r \in R} U(r) P_i(r) dr, \quad i = 1, 2, \dots, \quad (1)$$

where $U(r)$ is an utility function defined by preference patterns. Thus, to be consistent with rationality axioms, selection should maximise expected utility. To violate this rule is to be what SHARPE (1970) calls "irrational" or "unthinking". We are interested in finding out if a decision based on portfolio theory can be cast as a decision based on (1).

Many selection strategies in animal breeding are based on maximising the expectation of r . For such cases $U(r)$ is an identity operator. Because the goal of selection is to maximise r , we can argue objectively that $U(r)$ should be monotone increasing. However, there is no theory that forces $U(r)$ to be an identity operator. Some have argued implicitly that $U(r) = r$ due to infinite replication (e.g. BULMER, 1980). This argument is esoteric when decisions are not replicated. Consider a decision to make one of two choices, A or B. If we make A we are guaranteed \$ 1 000. However, if we make B we will do one of two things ; we might pay \$ 1 000 with probability $\frac{1}{2}$ or we might be paid \$ 3 000 with probability $\frac{1}{2}$. Most people would probably make the sure bet, choice A. There might be a few gamblers that take choice B. The quantitative geneticist who argues infinite replication is indifferent to the choices, because each has equal expected rewards. Do we really want $U(r) = r$?

Utility theory need not contradict classical selection theory. For example, selection strategies are sometimes invariant to the shape of $U(r)$ if $U(r)$ is at least monotone increasing and given equal uncertainty across all selection alternatives, i.e., equal accuracies such as exist with mass selection. Usually, it is when classical assumptions break down that we have contradictions. In general, $U(r)$ will be affected by an individual's gambling philosophy and this will affect selection decisions.

III. Portfolio theory

A. Description

Only a segment of portfolio theory is investigated here ; that indicated by the objective function described by SCHNEEBERGER *et al.* (1982).

With the approach of SCHNEEBERGER *et al.* (1982)

$$r = w s$$

where w is a row vector of non-negative weights which sum to one (or a number less than one) and s is a column vector of economic merits associated with sires. The vector w will define the usage of sires in the breeding program. Different realizations of w correspond to different selection alternatives. The goal of selection is to find w such that r is maximum under certain constraints.

Given w and information on sires it is possible to compute the mean and variance of r , as outlined by SCHNEEBERGER *et al.* (1982). Define

$$E_w(r) = M_w \text{ and } \text{Var}_w(r) = V_w.$$

The subscript w is a reminder that the mean and variance are functions of w . With the portfolio approach, w is chosen so as to maximise

$$M_w + k V_w, \quad (2)$$

where k is a constant that reflects an individual's attitude about risk. When k is negative, zero and positive the attitudes are described as risk averse, risk neutral and risk prone, respectively (SCHNEEBERGER *et al.*, 1982). Objective (2) is maximised subject to linear and non-negativity constraints.

B. Accusation

Given the importance of utility theory, it is natural to suspect that past attempts have been made to show the equivalence or non-equivalence of utility theory and maximising (2). Indeed, NAYLOR & VERNON (1969, p. 420) claim that (2) can be equated to expected utility. However, the proof which is found in FARRAR (1962, p. 20) seems questionable. The derivation of FARRAR is based on expanding $U(r)$ by a Taylor series and taking expectation. A similar approximation of expected utility is given by ANDERSON *et al.* (1977) and is

$$E_w[U(r)] \doteq U(M_w) + \frac{1}{2} U''(M_w) V_w \quad (3)$$

where $U''(M_w)$ is the second derivative of $U(r)$ evaluated at $r = M_w$. We see no way of equating (2) and the right side of (3). Moreover, it is easy to show that (2) cannot equal the expectation of some utility function when $k \neq 0$ (see Appendix).

SHARPE (1970) argues that an utility function must be quadratic if it is to be consistent with the concerns of those who would use portfolio theory, i.e. concerns about M_w and V_w . Though SHARPE recognizes the contributions of utility theory, he ultimately discards it because a quadratic utility function implies an unrealistic behavior. However, we do not agree that concerns only with M_w and V_w imply a quadratic utility function. For example, when r is a normal random variable expected utility is an exclusive

function of the mean and variance and this result is independent of the utility function. Thus, we regard the negative argument of SHARPE as an incomplete justification of portfolio theory.

If maximising (2) is inconsistent with utility theory one might wonder how (2) accommodates risk. If (2) is not expected utility what is it? A justification for the use of portfolio theory exists and it is consistent with utility theory.

C. Acquittal

We construct a rational decision rule that is equivalent to the portfolio approach. First we will design a monotone utility function, say $U^+(r)$, possessing and attribute which seems reasonable from a portfolio outlook. There is a constancy about risk as it is treated by (2) ; decisions based on (2) are invariant to location shifts of r . In fact, location invariance is a major property from which significant results follow. The utility function, $U^+(r)$, will be such that the resulting decision rule is invariant to location shifts.

The equation,

$$U^+(r) = U^+(r + h) P + U^+(r - h) (1 - P) \tag{4}$$

represents a statement of indifference. For any h and r a P ($P \in [0,1]$) can be found where (4) holds. Consider a decision to make one of two choices, A or B. If A is made we will obtain r with probability one. If B is made we could obtain $r + h$ with probability P or we could obtain $r - h$ with probability $1 - P$. Equation (4) says that values associated with gambles A and B are equal.

If the decision rule based on $U^+(r)$ is invariant to location shifts, then (4) must hold for all r because statements of indifference are decisions. Thus, for any h (4) holds for all r , where P is a function of h but not r . In the Appendix, we prove that $U^+(r)$ can only be a linear or exponential function.

When $U^+(r)$ is strictly monotone increasing

$$U^+(r) = \begin{cases} a - be^{-cr} & (5.1) \\ a + br & (5.2) \\ a + be^{cr} & (5.3) \end{cases}$$

$b > 0, c > 0$, depending on one's approach to risk. It can be shown that the utility functions given by (5.1), (5.2) and (5.3) produce decision rules which are invariant to location shifts. These decision rules are also invariant to the constants a and b .

PRATT (1964) has derived (5.1), (5.2) and (5.3) from consideration of a function (i.e., $U'(r)/U''(r)$) that depicts the degree of risk. PRATT referred to utility (5.1) as concave and corresponding to a decision-maker with a constant aversion to risk. Likewise, utilities (5.2) and (5.3) correspond to decision-makers with constant attitudes about risk being risk neutral and risk prone, respectively. Note also that the parameter c in (5.1) and (5.3) acts as a scale effect in quantifying risk.

The insistence that decisions should only be based on M_w and V_w , as in portfolio theory, suggests a strong preoccupation with normality. A distribution which possesses only the first two cumulants is, by necessity, normal. A stronger argument is given in the Appendix that supports the belief that r is treated as normal. Moreover, we will see later that when r is not normal, decisions based on (2) can be irrational. One of the shortcomings of previous use of portfolio theory is that distributional assumptions

are rarely made explicit. Normality seems to be an assumption, though SHARPE (1970) gives a contrary view. All assumptions implied or otherwise should be highlighted so that the feasibility of (2) can be assessed.

Assume that for any w , r has a normal distribution. If $U^+(r)$ equals (5.1), the expected utility is

$$a - b \exp [-cM_w + \frac{1}{2} c^2 V_w]. \quad (6)$$

Now maximising (6) is tantamount to maximising

$$M_w - \frac{1}{2} c V_w.$$

Likewise, maximising expectations of (5.2) and (5.3) are the same as maximising

$$M_w \text{ and } M_w + \frac{1}{2} c V_w,$$

respectively. We have clearly vindicated the proposal of SCHNEEBERGER *et al.* (1982).

A rational person involved in decision-making must act as if he had implemented Bayesian machinery (e.g. a utility function). However, such a person need not have carried out a Bayesian analysis (BERGER, 1980). That is, rational person could have arrived at a decision in an indirect way. The portfolio approach can be considered an example of an indirect process.

D. Ramifications

Having shown what maximising (2) implies, we can now ask the question ; are the hidden assumptions justified? The hidden assumptions are :

- 1) Normality.
- 2) The utility function is a given by (5.1), (5.2) or (5.3).

It is not possible to categorically refute or support these assumptions as applied to particular problems. We can only suggest what might happen when assumptions are not satisfied and comment on what alternative procedures might be used.

When normality breaks down decisions based on (2) can be irrational. For example, assume that k in (2) is -0.05 (risk aversion). Now consider making one of two decisions, A or B. If A is made a sure \$ 100 is earned. If B is made \$ 150 could be earned with probability $\frac{1}{2}$ or \$ 300 could be earned with probability $\frac{1}{2}$. Because variance is penalised, A will be picked over B if (2) is used as a criterion. Yet no rational person would ever pick A over B. Alternatively, if expected utility is used, B will always be picked over A regardless of the risk philosophy, i.e. if the correct distributions are used to evaluate expectations.

The practitioner may be unwilling to use one of the utility functions, (5.1), (5.2) or (5.3). Alternative utility functions can be found in ANDERSON *et al.* (1977) and PRATT (1964). Evaluating the expectation of a more general utility function could be done with techniques described in SMITH & ALLAIRE (1985). Approximation (3) might also prove to be useful.

There are other complications with the use of (2) that should be considered. Until now, r has been taken as a univariate quantity. However, in practice r will typically be multivariate. Decisions based on expected utility of multivariate r are at least conceptually clear. However, a multivariate extension of (2) is not so obvious. Compli-

cations also occur when intermediate r is optimum. For this case, an appropriate utility function is likely non-monotone and use of (2) is inappropriate. Whereas decisions based on (2) can be rational, the explicit use of utility theory and related Bayesian methods offers greater flexibility.

Pure risk averse attitudes (e.g. when k in (2) is negative) are sometimes doubtful when long-term objectives are sought. An example where this is true is gene pool selection. SMITH & ALLAIRE (1985) described the additive merit (a) of an individual taken at random from some gene pool as: $a = a_1 + a_2$, where a_1 equals the average additive genetic effect of all base animals that contribute genes to the pool and a_2 is the additive genetic effect due to segregation in the random mating gene pool. We will assume that it is desired to maximise long-term additive merit. In selecting "parents" of the pool, a_1 and a_2 should be considered separately. Practitioners are liable to have a risk averse attitude when considering the a_1 that would result from various selection alternatives. However, a risk averse attitude about a_2 is unreasonable. Concerns about long-term response imply that variance of a_2 is desirable. This variance will tend to decrease with increasing inbreeding (i.e., drift) in the gene pool. The reward of any selection alternative may be taken as the vector (a_1, v) , where v is the variance of a_2 . A selection alternative influences v as a function of the number of parents of each sex and the relationships among the parents. Utility functions of (a_1, v) that might find application and are risk averse with respect to a_1 are represented by: $U(a_1, v) = -\exp\{-ca_1 - f(v)\}$ where $c > 0$ and $f(\cdot)$ is monotone increasing. There may be some breeders who are risk prone with respect to a_1 . For them, utility functions of the form $U(a_1, v) = \exp\{ca_1 + f(v)\}$ may be useful. Breeders who are risk neutral with respect to a_1 might try $a_1 + f(v)$. These three classes represent all rational decision rules which are: unaffected by shifts in the genetic base; and consistent with maximising a_1 and v .

Consider the vector of weights, w , defined for the portfolio approach and let it represent the proportion of genes that various parents contribute to the gene pool. Given the assumption of panmixia, the genomic table described in SMITH & ALLAIRE (1985) can be manipulated to show that $v = \sigma^2(1 - \frac{1}{2}w'Aw)$, where σ^2 is the additive genetic variance and A is the relationship matrix for parents. If it is desired to calculate w so as to maximise expected utility where $U(a_1, v) = -\exp\{-ca_1 - dv\}$, $\exp\{ca_1 + dv\}$ or $a_1 + dv$, $d > 0$, then objective (2) can be used by adding $-bw'Aw$, where $b = (d\sigma^2)/(2c)$ or $d\sigma^2$. The modified objective function is still a quadratic function in w and hence quadratic programming can be implemented to find the optimum w . For other problems, selection alternatives may be associated with uncertain w . This situation is harder but we might take advantage of the relation $E[w'Aw] = \text{tr}[A(qq' + Q)]$, where $q = E[w]$ and $Q = \text{Var}(w)$.

In the next section we describe applications of utility theory with mate selection. In particular we describe a rational objective which is consistent with the underlying reason for positive assortative mating. Unlike concerns about risk, the philosophy behind positive assortative mating can be justified under conditions of infinite replication. Indeed, concerns about risk appear to be opposite to this philosophy.

IV. Mate selection

Mate selection is defined as the selection of parents and the formation of mating pairs (ALLAIRE, 1980; SMITH & ALLAIRE, 1985). ALLAIRE (1980) has suggested that non-random mating practices imply a non-linear objective.

Mate selection was originally proposed as a method to enhance the value of replacement animals where value is determined by a non-linear merit function. However, here we consider the improvement of additive genetic merit rather than with non-linear merit. The formation of special mating pairs will do little to enhance the average additive merit of the progeny produced by those matings. However, if the objective of selection is to enhance average merit after several generations, mate selection is a useful tool. We may infer that this objective is non-linear.

Selection and positive assortative mating is an example of a mate selection rule that can enhance long-term additive merit. Selection works directly to increase merit in the progeny. Position assortative mating works to increase heritability in the progeny and thus the merit in the second generation. McBRIDE & ROBERTSON (1963) have demonstrated that selection coupled with positive assortative mating can lead to a greater selection response than selection with random mating. Combined selection and assortative mating is beneficial when heritability is high, selection intensity is low and the trait is polygenic (DE LANGE, 1974). BAKER (1973) claimed that assortative mating will increase response to selection by no more than 4 or 5 p. 100 in most situations. This assessment is based on populations where the fraction selected is 20 p. 100 or less. We agree with BAKER but note his figures are underestimated by as much as two percentage units because selection response was assumed proportional to the genotypic standard deviation. This is only so when heritability is close to one. When heritability is close to zero, response is proportional to the genotypic variance. A final point of contention is that BAKER implicitly assumed normality in the offspring of selected parents. The appropriateness of this assumption needs to be investigated when heritability is high and when parents are mated assortatively.

Early studies on combined selection and assortative mating seem to imply that assortative mating is done to enhance selection response by increasing heritability in the progeny. For example, BAKER (1973) evaluated assortative mating under conditions of mass selection, which is reasonable if heritability is the primary concern. However, we suggest that there are good reasons to consider different selection schemes and in particular use of a selection index that incorporates information on preassorted relatives. Heritability needs not be the sole concern.

Let I_p denote a set of records on parents and let C specify the way parents are mated. For example, C indicates the mating pattern such as assortative mating or a realization from a random mating scenario. Let $C = C_a$ when parents are selected and mated assortatively on I_p . Then, conditional on $I_p \cup C_a$ where \cup is the union operator, all other aspects of assortative mating are redundant. On the average, assortative mating causes an increase in the additive genetic variance of progeny compared to cases when C denotes random mating. Indeed, assortative mating differentiates progeny so that some progeny are more desirable than others and the estimated rank order can be determined from $I_p \cup C_a$.

Let I_o denote a set of records on offspring of the parents previously described. To carry out mass selection on progeny we can use the information $I_o \cup C_a$. However, because of sampling the rank order estimated from $I_o \cup C_a$ will not be the same as the rank order estimated from $I_p \cup C_a$. As the rank orders become less alike the prior differentiation of progeny becomes more superfluous and at the limit assortative mating will show no advantage. The appropriate thing to do is to use all the information, i.e. $I_p \cup I_o \cup C_a$, because : the resulting selection index or selection rule will have maximum accuracy ; the rank orders determined from $I_p \cup I_o \cup C_a$ and $I_p \cup C_a$ will generally be closer than those determined by $I_o \cup C_a$ and $I_p \cup C_a$.

In effect, we have argued that assortative mating is implemented to : create extreme positive individuals ; arrange future pedigree information (in the attempt to create extreme positive individuals). Therefore, an evaluation of assortative mating should at least consider progeny selected by an index which includes information on all preassorted relatives. With such a selection scheme SMITH & HAMMOND (1987) showed that the relative efficiencies of assortative mating over random mating are significantly larger than when mass selection is used.

Mate selection, as a theoretical decision problem, can be very difficult. Concerns about mating pairs and additivity imply a long term perspective. Whilst evaluating expected utility or merit for the next generation is straightforward, extensions beyond one generation are complicated. On the other hand, the practice of selection and assortative mating seems to be relatively easy ; simply select and mate in accordance to rank order of estimated merit. However, there are two practical concerns that could cause problems for the intuitive approach, viz.

- (1) Realistic constraints on the number of times an animal might be used in mating (not used, once or several times) should be considered in a joint selection and mating problem. We refer to these constraints as mating constraints.
- (2) Additive merit is estimated with error and usually with unequal information. A rational procedure that assigns a value to a particular mating combination is desired and we cannot be sure that consideration of only estimated merits, whilst neglecting the errors of estimation, is enough.

We do not attempt to design a complicated mate selection rule by considering an extended planning horizon. Rather, we describe a simple mate selection rule that is consistent with the salient features of associative mating. To mimic selection and assortative mating, the Bayesian would define probability distributions for the additive merit of potential progeny and identify those distributions that are regarded as desirable. The more extreme and positive the progeny are likely to be the more desirable the distributions.

For mate selection formulations of JANSEN & WILTON (1985) and SMITH & ALLAIRE (1985), utility is assigned to individual mating combinations and summed over all combinations in the selection alternative. Consider an objective function of the form

$$\sum_{ij} x_{ij} E\{F(A_{ij})\} \tag{7}$$

where $x_{ij} = 0$ or 1 and A_{ij} is additive merit (perhaps aggregate) for the progeny produced by mating the i -th male, to the j -th female (or egg cell). The mate selection problem is solved by finding those x_{ij} such that (7) is maximum subject to mating constraints. If $x_{ij} = 1$ when the solution is found the i -th male is to be mated with the j -th female.

In (7), $f(A)$ is understood to be monotone increasing and thus consistent with maximising additive merit. Because objective (7) represents a simple sum of contributions from would be progeny, it is not good at quantifying risk (and in particular the risk due to drift). Hence, it is necessary to place constraints (perhaps irrational) on x_{ij} such as

$$\sum_j x_{ij} \leq n_i$$

where n_i is the maximum number of times the i -th male can be mated ; and

$$\sum_i x_{ij} \leq 1 \text{ for all } j$$

$$\sum_{ij} x_{ij} = N$$

where N is the number of progeny to be selected. The top constraints are needed as it is typical for females to be mated once or not at all. It may be desired to avoid certain mating combinations (denoted by the index set Ω) because of inbreeding. These types of constraints are enforced passively by subtracting a sufficiently large cost from (7) when $x_{ij} = 1$ and $i, j \in \Omega$.

The mate selection problem can be solved by linear programming techniques (JANSEN & WILTON, 1985) and problem (1) listed above can be accommodated.

Problem (2) can also be dealt with. Our development on previous mate selection theory is the incorporation of $f(\cdot)$. If $f(\cdot)$ is monotone and convex, objective (7) places a rational premium on extreme positive individuals ; mate selection is coupled with positive assortative mating. If $f(\cdot)$ is monotone and concave mate selection induces the negative assortative mating of selected parents. If $f(\cdot)$ is linear, no particular mating combination is favoured within some feasible set. In addition to these properties, the strategy will assign a value to the variability of A_{ij} depending on the curvature of $f(\cdot)$. If $f(\cdot)$ is convex, then variability is favoured and this is desirable from a long-term perspective (VAN RADEN *et al.* (1984) made a similar argument in a sire problem).

What kind of function should $f(\cdot)$ be? At the very least it should be monotone increasing and if positive assortative mating is desired, convex. Numerous functions fit these requirements but there is only one function, (5.3), that results in a decision rule unaffected by location shifts (objective (7) can usually be represented as $E[f(r)]$ where r is a randomized reward). When the additive genetic mean is not well defined, use of (5.3) is proposed. The objective function (7) would then be a sum of expectation of functions like (5.3).

We give a brief illustration. Assume equal information so that variance terms can be ignored and let $f(\cdot) = \exp(\cdot)$. It is desired to mate two males (M_1, M_2) to two females (F_1, F_2) so that all animals are mated once and objective (7) is maximum. There are two selection alternatives : ($M_1 \times F_1, M_2 \times F_2$) and ($M_1 \times F_2, M_2 \times F_1$). Let the estimated merits be 1.1, 2.0, 1.5, 1.9 for M_1, M_2, F_1 and F_2 , respectively. Under conditions of normality, (7) is proportional to

$$\exp \{ \frac{1}{2} (1.1 + 1.5) \} + \exp \{ \frac{1}{2} (2.0 + 1.9) \} = 10.7$$

for ($M_1 \times F_1, M_2 \times F_2$) and

$$\exp \{ \frac{1}{2} (1.1 + 1.9) \} + \exp \{ \frac{1}{2} (2.0 + 1.5) \} = 10.2$$

for ($M_1 \times F_2, M_2 \times F_1$). Objective (7) is maximum when parents are mated assortatively.

V. Discussion

We have investigated a component of portfolio theory and found it to be rational provided certain assumptions are satisfied. As a side issue we described an interesting way of viewing mate selection. Consideration of rationality, risk and information has proven very useful.

In our formulation of mate selection, we have assumed that total utility can be represented as a sum of utilities from individuals in the selected population. Moreover, relationships among animals and concerns about risk have been ignored in the utility function. This is tantamount to assuming that the selected population comprises an infinite number of sufficiently unrelated animals. Lastly, we have assumed that loading additive merits into extreme positive individuals is desirable. This is reasonable when information is reused from generation to generation. In a multiple-trait setting, the reuse of information may be critical because estimated merit is affected not just by economic weights but by a complicated covariance structure. Mating parents assortatively by an economic index will almost surely be suboptimal when information is not reused. A different criterion of desirability one based on maximising accuracy of selection, was given by FERNANDO & GIANOLA (1984) for the multiple-trait case. However, these authors did not consider selection and mating together.

It is difficult to make general statements about mate selection using objective (7). However, the proposal is amenable to simulation under various constraints (e.g. population size and structure, level of information, selection pressure) and values of c in (5.3).

Objective (7) can be evaluated by use of mixed model methodology, as outlined by SMITH & ALLAIRE (1985). Hence mate selection rules can accommodate unequal information, selection or mating bias, fixed effects, etc. More importantly, mixed model methodology allows the reuse of information in mate selection problems.

Received July 18, 1986.

Accepted January 26, 1987.

Acknowledgements

The authors wish to thank Dr R. BARLOW, Mr. P.F. PARNELL (both with the NSW Department of Agriculture, Australia) and the reviewers for their useful remarks.

References

- ALLAIRE F.R., 1980. Mate selection by selection index theory. *Theor. Appl. Genet.*, **57**, 267-272.
ANDERSON J.R., DILLON J.L., HARDAKER B., 1977. *Agricultural decision analysis*. 344 p., The Iowa State University Press, Ames, Iowa.
BAKER R.J., 1973. Assortative mating and artificial selection. *Heredity*, **31**, 231-238.
BERGER J.O., 1980. *Statistical decision theory*. 425 p., Springer-Verlag, New York, Inc.

- BULMER M.G., 1980. *The mathematical theory of quantitative genetics*. 255 p., Oxford University Press, Walton Street, Oxford.
- DE LANGE A.D., 1974. A simulation study of the effects of assortative mating on the response to selection. *1st World Congr. Genet. App. Livestock Prod., Madrid, 7-11 October 1974*, **3**, 421-425, Editorial Garsi, Madrid.
- DILLON J.L., 1977. *The analysis of response in crop and livestock production*. 2nd ed., 213 p., Pergamon Press, Oxford.
- FARRAR D.E., 1962. *The investment decision under uncertainty*. 90 p., Prentice Hall, Inc, Englewood Cliffs, N.J.
- FERNANDO R.L., GIANOLA D., 1984. Rules for assortative mating in relation to selection for linear merit functions. *Theor. Appl. Genet.*, **68**, 227-237.
- JANSEN G.B., WILTON J.W., 1985. Selecting mating pairs with linear programming techniques. *J. Dairy Sci.*, **68**, 1302-1305.
- MCBRIDE G., ROBERTSON A., 1963. Selection using assortative mating in *D. melanogaster*. *Genet. Res.*, **4**, 356-369.
- NAYLOR T.H., VERNON J.M., 1969. *Microeconomics and decision models for the firm*. 482 p., Harcourt, Brace and world, New York.
- PRATT J.W., 1964. Risk aversion in the small and in the large. *Econometrica*, **32**, 122-136.
- ROBERTSON A., 1960. A theory of limits in artificial selection. *Proceedings of the Royal Society of London*, B, 153, 234-249.
- SCHNEEBERGER M., FREEMAN A.E., BOEHLJE M.D., 1982. Application of portfolio theory to dairy sire selection. *J. Dairy Sci.*, **65**, 404-409.
- SHARPE W.F., 1970. *Portfolio theory and capital markets*. 316 p., McGraw Hill, Inc, New York.
- SMITH C., 1985. Scope for selecting many breeding stocks of possible economic value in the future. *Anim. Prod.*, **41**, 403-412.
- SMITH S.P., ALLAIRE F.R., 1985. Efficient selection rules to increase non-linear merit : application in mate selection. *Génét. Sél. Evol.*, **17**, 387-406.
- SMITH S.P., HAMMOND K., 1987. Assortative mating and artificial selection : a second appraisal. *Génét. Sél. Evol.*, **19**, 181-196.
- VAN RADEN P.M., FREEMAN A.E., ROTHSCHILD M.F., 1984. Maximising genetic gain under multiple-stage selection. *J. Dairy Sci.*, **67**, 1761-1766.

Appendix

(a) Proof that objective (2) cannot equal expected utility

To show that (2) cannot equal expected utility we make an argument of contradiction. Assume there exists a utility function, $U^*(r)$, such that

$$E_p[U^*(r)] = M_p + k V_p.$$

where the subscript p, in this case, signifies a probability distribution function for r. Consider a hypothetical sequence of probability distribution functions, say P_1, P_2, \dots , such that

$$\lim_{n \rightarrow \infty} M_{p_n} = x \text{ and}$$

$$\lim_{n \rightarrow \infty} V_{p_n} = 0$$

where the subscript p_n represents an indexed distribution. We are assuming that the set of allowable distributions [i.e., the set of distribution for which (2) applies] is rich

enough to include such a sequence for any real number x . If the set is not rich enough then the use of (2) would be severely restricted and nothing else need be said. The set of normal distributions is rich enough to include (P_1, P_2, \dots) . It follows that

$$\lim_{n \rightarrow \infty} E_{P_n} [U^*(r)] = x$$

The limiting distribution, P_∞ , can be thought of as a function that assigns a probability mass of one to $r = x$. Thus, with appropriate regularity conditions, the right side of (8) can be equated to $U^*(x)$. Hence, we have shown that

$$U^*(x) = x$$

and this is true for all x (in our formulation x is unspecified). However, for a general distribution,

$$E_p[U^*(r)] = M_p \neq M_p + k V_p$$

when $k \neq 0$. Thus, we have arrived at a contradiction.

(b) Proof that $U^+(r)$ is a linear or exponential function

First consider the case where $P \neq 0$ or 1 in (4). Equation (4) can be written as

$$V(r + h, h) = G(h)V(r, h) \tag{9}$$

where

$$G(h) = \frac{1 - P}{P}$$

and

$$V(r, h) = U^+(r) - U^+(r - h).$$

Like (4), (9) holds for all r . Consequently, recursive application of (9) yields :

$$V(r + nh, h) = G(h)^n V(r, h). \tag{10}$$

Now by definition

$$\sum_{n=1}^N V(R + nh, h) = U^+(r + Nh) - U^+(r).$$

Applying the summation to the right side of (10) gives :

$$\sum_{n=1}^N G(h)^n V(r, h) = \begin{cases} N V(r, h) & \text{if } G(h) = 1 \\ \frac{G(h) - G(h)^{N+1}}{1 - G(h)} V(r, h), & \text{otherwise} \end{cases}$$

Thus, we have shown that $U^+(r + Nh)$ equals

$$NV(r, h) + U^+(r) \tag{11.1}$$

or

$$\frac{G(h) - G(h)^{N+1}}{1 - G(h)} V(r, h) + U^+(r) \tag{11.2}$$

depending on $G(h)$. The derivation assumes N is a positive integer. However, identities (11.1) and (11.2) hold when N is a non-positive integer as well. When $N = 0$ this is obvious. To see the validity of (11.1) and (11.2) when N is negative, note that (9) implies

$$V(r - h, h) = G(h)^{-1} V(r, h) \tag{12}$$

Recursive application of (12) yields

$$V(r - nh, h) = G(h)^{-n} V(r, h) \tag{13}$$

Summing both sides of (13) from $n = 0$ to $n = N - 1$ leads to identities (11.1) and (11.2) where N is allowed to be negative. Furthermore, if $P = 0$ or 1 , it is easy to show that (11.1) holds where $V(r, h) = 0$.

Taking $r = 0$, we see that $U^+(Nh)$ is either a linear or exponential (with intercept) function of the whole number N . Alternatively, $U^+(t)$ is a linear or exponential function (say $g(t)$) if $t \in \{Nh : N = 0, \pm 1, \pm 2, \dots\} = \Omega(h)$. When $t \in \Omega(Nh^* : N = 0, \pm 1, \pm 2, \dots) = \Omega(h^*)$ and $h^* \neq h$, $U^+(t)$ is also a linear or exponential function (say $g^*(t)$). However, we cannot be sure that $g^*(t) = g(t)$. If $h^* = Kh$, where K is a rational number, then there are an infinite number of points common to both $\Omega(h)$ and $\Omega(h^*)$. In which case $g^*(t) = g(t)$ because $U^+(t) = g(t) = g^*(t)$ if $t \in \Omega(h) \cap \Omega(h^*)$ where \cap is the intersection operator. This implies $U^+(t) = g(t)$ if $t \in \{Kh : K \text{ rational}\} = \Omega$. Assuming that $U^+(z) \neq g(z)$ for some $z \notin \Omega$, leads to a contradiction. All neighbourhoods of z contain points contained in Ω . Monotonicity of $U^+(\cdot)$ implies

$$g(z_1) \leq U^+(z) \leq g(z_2) \tag{14}$$

where $z_1 < z < z_2$ and $z_1, z_2 \in \Omega$.

The upper and lower bounds in (14) can be made sharper and sharper thus inducing the desired contradiction. This shows that $U^+(t) = g(t)$ for all t .

(c) Strong argument that suggests objective (2) implies normality

To suggest that r is treated as normal we will assume that use of objective (2) is consistent with maximising expected utility. If this is not true, the portfolio approach would be irrational and nothing else need be said. As objective (2) is unaffected by location shifts the underlying utility function is almost certainly one of (5.1), (5.2) or (5.3). Note first that when the utility function is linear [i.e., (5.2)] normality need not hold; expected utility involves only the mean and this result is independent of the distribution. Thus, the distribution of r is only an issue when the utility function is exponential [i.e., (5.1) or (5.3)].

If maximising objective (2) is equivalent to maximising expected utility (EU), then there must be some function of EU, a , b and c [say $f(\text{EU}, a, b, c)$] that equals objective (2) [recall a , b and c come from (5.1) or (5.3)]. The functional form of f is independent of the distribution of r . However, it must be that maximising f is equivalent to maximising EU. To proceed further we must determine f .

As with part a) of the Appendix we will use a sequence of probability distribution functions, say P_1, P_2, \dots , such that

$$\lim_{n \rightarrow \infty} M_{P_n} = x \text{ and}$$

$$\lim_{n \rightarrow \infty} V_{P_n} = 0$$

We assume that the set of allowable distributions is rich enough to include such a sequence for any real number x . Moreover, we assume that f is a continuous function of EU . When the utility function equals (5.3) [a similar argument can be made when utility equals (5.1)] the limiting value of EU is

$$a + be^{cx}$$

The limiting value of f or objective (2) is x . It follows that f is defined implicitly for all x by

$$f(a + be^{cx}, a, b, c) = x$$

and this implies

$$f(EU, a, b, c) = \frac{1}{2} \log [(EU - a)b]$$

Note that $(EU - a)/b$ is the moment generating function {i.e., $E[\exp(cr)]$ } for r . Equating f to (2) we obtain

$$E[\exp(cr)] = \exp [cM + \frac{1}{2}ckV] \tag{15}$$

where M and V are the mean and variance of r , respectively. To prove normality it suffices to show that $k = \frac{1}{2}c$. We know only that k is a function of c [say $k(c)$].

We will now assume that if r is a member of the class of allowable distributions so is tr where t is any positive constant. The mean and variance of tr are

$$tM \text{ and } t^2V.$$

Thus, the moment generating function given by (15) is

$$E[\exp(ctr)] = \exp [ctM + \frac{1}{2}ckt^2V]. \tag{16}$$

If we define $c^* = ct$, from (15) we obtain

$$E[\exp(c^*r)] = \exp [c^*M + \frac{1}{2}c^*k^*V] \tag{17}$$

where $k^* = k(ct)$. Equating (16) and (17) shows that

$$ck(c)t^2 = ctk(ct)$$

and setting $c = 1$ yields $k(t) = k(1)t$. That is, $k(t)$ [or $k(c)$] is proportional to t (or c). As the log of (15) must be a legitimate cumulant generating function it follows that $k(c) = \frac{1}{2}c$.