THE RAMIFICATIONS OF PORTFOLIO AND UTILITY THEORIES FOR DAIRY BREEDING PROGRAMS

S.P. SMITH and K. HAMMOND, AUSTRALIA

Animal Genetics and Breeding Unit, University of New England, Armidale NSW 2350 Australia.

SUMMARY

Decision rules based on maximising expected utility can be justified by strong Bayesian arguments. Such rules correctly accommodate both objective and subjective criteria. Portfolio theory, as applied to dairy sire selection problems, is consistent with maximising expected utility. However, the utility function is restricted and distributional assumptions must hold. A more direct approach is the explicit use of utility theory and related Bayesian methods. New applications for utility theory in animal selection are described. Two situations for long-term improvement (after several selection decisions) of additive merit are considered: gene pool selection and mate selection. Utility functions are given which accommodate subjective concerns about risk and genetic variability.

INTRODUCTION

Schneeberger et al (1982) proposed a portfolio theory as an aid to decision making in dairy cattle sire selection. This theory allows the joint quantification of risk and objective measures of merit. In this paper, Schneeberger's portfolio theory is recognized as a decision rule that gives maximum expected utility. The utility function is investigated and we mention how this and other utility functions can be used in breeding programs.

BAYESIAN DECISION THEORY

Bayesian decision theory, as applied to selection problems, is described by Smith and Allaire (1985). The objective of selection is to secure a desirable state of nature or reward, r. Here we will initially be concerned with a univariate r. Later we allow r to be an aggregate or a vector. The Bayes decision rule is the action (selection alternative), a, in action space, A, where

\[ E(U(r[a])|y) \]

is maximum (Berger, 1980). In (1), r[a] is the reward resulting from action a, U(.) is the utility function and y is a data vector.

There are strong justifications for use of a Bayes rule. First, note that (1) is a conditional expectation given data. Thus decisions are made from realized data, consequently pre-experiment views (e.g. those resulting from repetitive sampling) do not corrupt the decision rule. Second, if some decision rule is to be consistent with axioms of rational behaviour then it must be equivalent to some Bayes rule (Berger, 1980).
PORTFOLIO THEORY

With the approach of Schneeberger et al (1982)

\[ r = p s \]

where \( p \) is a row vector of non-negative weights and \( s \) is a column vector of merits (e.g. estimated breeding values corrected for semen costs) associated with sires. The vector, \( p \), defines the relative usage of the sires in the breeding program; the optimum \( p \) being dependent on the estimated breeding values of the set of sires, semen costs and attitude towards risk. Different realizations of \( p \) correspond to different actions or selection alternatives. A decision is needed to find \( p \) so as to increase and hopefully maximize \( r \).

Define

\[ E\{r|y\} = M[p] \text{ and } Var\{r|y\} = V[p] \]

where the bracketed \( p \) is a reminder that the mean and variance is a function of \( p \). \( M[p] \) and \( V[p] \) can be evaluated as outlined by Schneeberger et al (1982). We have taken liberties by defining \( M[p] \) and \( V[p] \) as the conditional moments given \( y \).

With portfolio theory, \( p \) is found such that

\[ M[p] + k V[p] \]

is maximum subject to user defined constraints. When \( k \) is negative, (2) penalizes variance and represents a decision-maker who is risk averse. When \( k \) is zero, (2) represents a decision-maker who is risk neutral. When \( k \) is positive, the decision-maker is risk prone.

Decisions based on (2) are unaffected by changes in the mean value of \( r \). When decisions based on (1) are unaffected by changes in mean, then Smith and Hammond (1985a) proved that \( U(r) \) must be a linear or exponential function. In particularly \( U(r) = r - \exp(-cr) \) or \( \exp(cr) \) where \( c > 0 \).

Concerns which deal only with \( M[p] \) and \( V[p] \), as in portfolio theory, suggest that the posterior distribution of \( r \) given \( y \) is normal. Smith and Hammond (1985a) showed that decisions based on (1) are equivalent to those based on (2) when \( U(r) \) is a linear or exponential function (\(|k| = \frac{1}{2} c\)) and when \( r \) is normal. This shows that decisions based on (2) are rational, given these conditions.

For most cases where the improvement realised is for one stage of selection only, \( k \) is negative or equivalently \( U(r) = -\exp(-cr) \). Pratt (1964) referred to this utility function as concave and corresponding to a decision-maker with a constant aversion to risk (i.e. decisions are unaffected by changes in the mean).

RAMIFICATIONS

Whereas decisions based on (2) might be rational, care is required in the use of (2). When normality does not hold, decisions based on (2) can be irrational (Smith and Hammond, 1985a). Moreover, the utility function of
interest may not be linear or exponential. Use of objective (1) is safer because it allows for non-normality and various utility functions.

The utility function can take on various forms depending upon how and why it's used. Even if a unique objective exists, the utility function may change according to whether short-term (after one selection decision) or long-term improvements are sought. Consider a situation where \( r \) represents an average breeding value for a closed population and assume an intermediate reward (\( r^* \)) is desired, e.g. miling speed. Goddard (1983) proposes a linear selection index which is consistent with the long-term objective of moving \( r \) closer to \( r^* \). Any utility function of the form \( U(r) = g(r-r^*) \), where \( g(.) \) is unimodal and maximum at the origin, is also consistent with moving \( r \) closer to \( r^* \). Moreover, \( U(r) \) may only have vague similarities with the merit function used to define \( r^* \). Alternatively, a utility function that resembles a merit function is sometimes reasonable when short-term improvements are sought. Such is the case of mate selection where reward is represented as a vector of phenotypes for any individual (Smith and Allaire, 1985). An utility function should be defined by the user so as to induce a response consistent with the objective and in line with the selection and/or mating strategy.

Pure risk aversive attitudes (e.g. \( k \) in (2) negative) are sometimes doubtful when long-term objectives are sought. We consider two examples, gene pool selection and mate selection, where pure risk avoidance is inappropriate.

Smith and 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 aversive attitude when considering the \( a[1] \) that would result from various selection alternatives. However, a risk aversive 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 \( a[2] \). Whereas, \( v \) is affected by the selection alternative, it can sometimes be regarded as constant given the selection alternative. Utility functions of \((a[1],v)\) that might find application are represented by: \( U(a[1],v) = -\exp(-ca[1]+f(v)) \) where \( c>0 \) and \( f(.) \) is monotone increasing. This class of utility functions forms a unique representation of decision rules which are; unaffected by shifts in the genetic base, risk aversive with respect to \( a[1] \), and consistent with maximising \( a[1] \) and \( v \). \( c \) and \( f(.) \) are determined subjectively.

There may be some breeders who are risk prone with respect to \( a[1] \). For them, utility functions of the form \( U(a[1],v) = \exp(ca[1]+f(v)) \), where \( c>0 \) and \( f(.) \) is monotone increasing, may be useful. Breeders who are risk neutral with respect to \( a[1] \) might try \( a[1]+f(v) \). These two classes represent all rational decision rules which are; unaffected by shifts in the genetic base, risk prone or risk neutral with respect to \( a[1] \), and consistent with maximising \( a[1] \) and \( v \).

For most mate selection formulations, utility is assigned to individual mating combinations and summed over all combinations in the selection.
alternative. Consider an objective function of the form
\[
\sum_{ij} x[i,j] E(U[r[i,j]])y
\] (3)
where \( x[i,j] = 0 \) or \( 1 \) and \( r[i,j] \) 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 \( x[i,j] \) such that (3) is maximum subject to constraints. If \( x[i,j] = 1 \) when the solution is found the \( i \)-th male is to be mated with the \( j \)-th female.

In (3), \( U(r) \) is understood to be monotone increasing and thus consistent with maximising additive merit. Objective (3) does not seem to be good at quantifying risk (and in particular the risk due to drift). Hence, it is necessary to place constraints (perhaps irrational) on \( x[i,j] \) such as
\[
\sum_{j} x[i,j] \leq n[i]
\]
where \( n[i] \) is the maximum number of times the \( i \)-th male can be mated. Other usage constraints are needed; viz.
\[
\sum_{i} x[i,j] \leq 1 \quad \text{for all } j
\]
\[
\sum_{ij} x[i,j] = N
\]
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 \( \Omega \)) because of inbreeding. These types of constraints are enforced passively by subtracting a sufficiently large cost from (3) when \( x[i,j] = 1 \) and \( i,j \in \Omega \). The mate selection problem can be solved by linear programming techniques (Jansen and Wilton, 1985).

When \( U(r) \) in (3) is monotone and convex, mate selection induces both selection and positive assortative mating. Selection and positive assortative mating has been shown to yield a larger response than selection and random mating (McBride and Robertson, 1963). Assortative mating is most beneficial when heritability is high, selection intensity is low and when the trait is polygenic (De Large, 1974). Baker (1973) claims that assortative mating will increase selection response but only by a small amount. However, Baker's projections are based under conditions of mass selection. Mate selection uses information and places a rational premium on extreme positive individuals. Information that was used to load extreme individuals should be reused to make selection decisions in subsequent generations. To do otherwise seems to contradict the purpose of loading. Smith and Hammond (1985b) showed that the relative efficiency of prior assortative mating is greatly enhanced under selection by an index that incorporates information on all preassorted relatives.

Like decisions based on (2), those based on (3) are unaffected by shifts in the genetic base when \( U(r) \) is a linear or exponential function. Hence, using \( U(r) = \exp(cr) \) (where \( c > 0 \)) is proposed when the genetic base is undefined and when positive assortative mating is desired.
Objective (3) can be evaluated by use of mixed model methodology as outlined by Smith and 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.

REFERENCES


