Mathematical Social Sciences 23 (1992) 255-282
North-Holland

Fundamental derivations from decision field theory

Jerome R. Busemeyer

Purdue University, West Lafayette, IN 47907, USA

James T. Townsend

Indiana University, Bloomington, IN 47405, USA

Communicated by A.A.J. Marley
Received 10 August 1990
Revised 27 June 1991

Decision field theory is a stochastic dynamic model of decision-making based on psychological principles of approach-avoidance behavior. This paper provides a summary of the main mathematical derivations for the distribution of choice response times obtained from binary choice tasks, and the distribution of matched values obtained from a matching task. The relations of decision field theory to a number of other theories of decision-making are also pointed out.

Key words: Decision-making; choice models; response time models; diffusion models.

1. Introduction

Decision field theory is a dynamic stochastic model of decision-making based on psychological principles of approach-avoidance behavior; similar principles were originally described by Lewin (1935), Hull (1938), Miller (1944), Indow (1958), Bower (1959), Estes (1960), Atkinson and Birch (1970) and, most recently, Coombs and Avrunin (1977). In three previous articles, we described applications of the theory to approach-avoidance behavior (Townsend and Busemeyer, 1989), choice between risky and uncertain courses of action (Busemeyer and Townsend, 1989), and matching judgments (Busemeyer and Goldstein, in press). One unique feature of this model is that it provides a unified theoretical treatment of a wide range of measures of preference including approach-avoidance movements, choice probability, choice response time, selling prices, buying prices, indifference judgments, and strength of preference ratings.

Our theory differs from most mathematical approaches to decision-making by being dynamic rather than static. The theory's dynamics emerge naturally from psychological aspects of the underlying process theory (Townsend and Busemeyer,

1989). The same may be said for the stochastic characteristics which follow from associating reasonable probabilistic notions with the foundational deterministic difference and differential equations (Busemeyer and Townsend, 1989; Busemeyer and Goldstein, in press). We also see a merging of theoretical concepts normally identified with other areas of psychology, such as memory search, psychological discrimination, and reaction time, with those from the choice and decision literature. The reader is referred to Link and Heath (1975), Ratcliff (1978), Townsend and Ashby (1983), and Luce (1986) for more background on those other spheres of discourse.

The purpose of this paper is not to present a broad overview of decision field theory and the approach-avoidance principles upon which it is based; instead, our purpose is to provide a fairly complete and detailed mathematical development of an important but much more limited part of this general theory. More specifically, we derive equations for choice probability and the distribution of choice response times obtained from the stochastic choice model that forms the nucleus of decision field theory. In addition, we derive equations for the distribution of matched values obtained from a stochastic matching model that is essential for linking choice to many other measures of preference strength (e.g. cash difference judgments). We conclude this paper by showing how a number of other models of choice are related to decision field theory, including von Neumann and Morgenstern's (1947) expected utility model, Tversky's (1969) additive difference model, Thurstone's (1959) random utility model, and Ratcliff's (1978) resonance model. The first half of the paper describes a discrete time, discrete state Markov chain approach to modeling choice, and the second half describes a continuous time, continuous state diffusion approach.

2. Discrete time and space Markov chain model

2.1. Basic assumptions

Suppose that the decision-maker is given a choice between two courses of action, where each course of action is defined by a set of possible consequences conditioned on events. For concreteness, suppose the choice is made on a computer by moving an index finger from a center button to either a left or a right button. Figure 1 provides an illustration of the proposed decision process.

The presentation of the choice stimulus evokes two processes—an excitatory and an inhibitory process. The excitatory process is represented by a variable, P, called the preference state, which represents the decision-maker's tendency to move toward each alternative.[1] Negative values of P represent a tendency to move

[1] The notation used in this paper was chosen to be consistent with the notation used in previous papers on decision field theory. We thank A. Diederich, R. Heath, and P. Smith for comments.

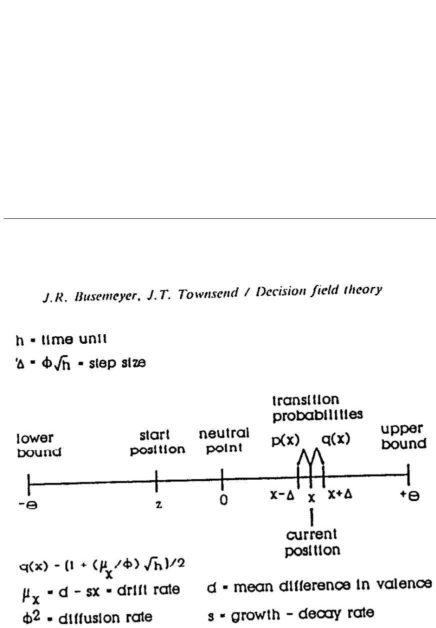

Fig. 1.

toward the left button, and positive values of P represent a tendency to move toward the right button. Immediately after the onset of the choice display, an initial state of preference is evoked, denoted by $P(0)=z$. This preference state changes and evolves as the decision-maker deliberates over the various possible consequences produced by each course of action, producing a new preference state at each moment in time, denoted $P(t)$.

The inhibitory process produces a tendency that opposes action. In general, the strength of the inhibitory process may gradually weaken or decay during deliberation as a function of factors such as the cost of waiting (see, for example, Busemeyer and Rapoport, 1988). However, for the purpose of this paper, we will assume the deliberation period is relatively short so that such cost factors are relatively constant, and so the strength of the inhibitory process remains constant during the deliberation period. The strength of this constant inhibitory tendency, represented by the magnitude, θ, is called the inhibitory bound (i.e. a response threshold). The left alternative is chosen as soon as $P(t) \le -\theta$ and the right alternative is chosen as soon as $P(t) \ge +\theta$. No physical movement takes place until the magnitude of the preference state overcomes the inhibitory bound, i.e. until $|P| \ge \theta$.

Preference is assumed to change by very small amounts during small time intervals, producing a gradual drift in preference over time. The symbol $h>0$ will denote a very small time interval. The change in preference during the time interval h is denoted $\mathrm{d}P(t)=P(t)-P(t-h)$. More specifically, $\mathrm{d}P(t)$ may be a very small positive step equal to $+\Delta=+\phi\cdot\sqrt{h}$, or a very small negative step equal to $-\Delta=-\phi\cdot\sqrt{h}$. The parameter Δ is called the step size, and the parameter ϕ^2 is called the diffusion rate for the finite state and time model. A continuous state and time model with infinitesimal time intervals and step sizes (described later), is obtained by letting the finite time interval, h, approach zero in the limit. Two basic assumptions are made about the transition probabilities:

Assumption 1: Time homogeneous, Markov transition probabilities. First, we assume that the following Markov property holds:

$$\Pr[\mathrm{d}P(t)=-\Delta \mid P(t-h)=x,\ P(t-2h)=y,\ldots]$$
$$=\Pr[\mathrm{d}P(t)=-\Delta \mid P(t-h)=x]=p(x),$$
$$\Pr[\mathrm{d}P(t)=+\Delta \mid P(t-h)=x,\ P(t-2h)=y,\ldots]$$
$$=\Pr[\mathrm{d}P(t)=+\Delta \mid P(t-h)=x]=q(x),$$

where p and q are functions that map the previous state of preference, $P(t-h)=x$, into the closed interval $[0,1]$ and $p(x)+q(x)=1$.

We could allow $r(x)=p(x)+q(x)<1$, or in other words we could allow steps of size zero without causing complications in the derivations that follow. However, this adds an additional parameter, i.e. the probability of the zero step, which we wish to avoid.

The transition probability functions $p(x)$ and $q(x)$ are spatially non-homogeneous since they depend on the state x. This deviates from previous choice models which assume transition densities that are independent of the state (Laming, 1968; Link and Heath, 1975; Ratcliff, 1978). Fortunately, this added flexibility comes at little cost in terms of number of parameters or complexity of the analyses.

The transition probability functions $p(x)$ and $q(x)$ are time homogenous since they do not depend on the time index. While it is possible to develop time non-homogeneous models (see Heath, 1992; Smith, 1990), this additional flexibility does increase the complexity of the analyses. In some (but not all) applications, this additional complexity may be required. However, even in the latter case, it is possible to use transition probabilities that are time homogeneous within large time intervals (Diederich, 1991).

The finite drift rate, denoted $\mu(x)$, and the finite diffusion rate, denoted ϕ^2, can be derived from the first and second raw moments of $\mathrm{d}P(t)$, conditioned on the previous state of preference, as follows:

$$\mu(x)=E[\mathrm{d}P(t) \mid P(t-h)=x]/h=[q(x)-p(x)]\cdot\Delta/h,$$
$$\phi^2=E[\mathrm{d}P(t)^2 \mid P(t-h)=x]/h=[q(x)+p(x)]\cdot\Delta^2/h.$$

Although not exhibited in $\mu(x)$ and ϕ^2 themselves, the finite drift and diffusion rates are also clearly functions of the finite time unit h, which is assumed to be a fixed constant. Its value is chosen a priori to be as close to zero as needed to make the discrete time Markov chain model approximate the continuous time, continuous state diffusion process as accurately as desired.

The transition probabilities $p(x)$ and $q(x)$ can be written as functions of the finite drift and diffusion rates by solving for $p(x)$ and $q(x)$ in terms of the parameters $(\mu(x),\phi^2,h)$ in the above two equations:[2]

[2] Occasionally we will substitute μ_x for $\mu(x)$. The former is more compact, which is useful in complex expressions. The latter is useful for emphasizing that the mean drift rate is a function of the position. However, they have equivalent meanings so that $\mu_x=\mu(x)$.

$$p(x) = [1 - (\mu_x/\phi) \cdot \sqrt{h}]/2, \tag{1a}$$

$$q(x) = [1 + (\mu_x/\phi) \cdot \sqrt{h}]/2. \tag{1b}$$

The above equations require that the following inequalities are satisfied in order to contain $0 \le p(x)$ and $q(x) \le 1$:

$$-1/\sqrt{h} \le (\mu_x/\phi) \le +1/\sqrt{h}.$$

Assumption 2: Linear Growth Model. The second basic assumption is that the drift rate is linearly related to the current state of preference:[3]

$$\mu(x) = d - s \cdot x. \tag{2a}$$

The parameter d is called the mean difference in valence. It represents the mean input in valence that drives the preference system over time. The parameter s is called the growth rate parameter. It determines the rate of growth toward an equilibrium point when the mean input is non-zero, and it represents the rate of decay toward zero (the neutral or rest point) when the mean input is zero (absent).

One special case of (2a) is called the proportional change growth model. It is obtained by assuming that $d = s \cdot d^*$ for some value d^*. In this case, $\mu(x) = s \cdot (d^* - x)$.

Another special case of (2a) is called the positive linear growth model. It is obtained by assuming that $s < (d/\theta)$ so that $\mu(x) > 0$ whenever $d > 0$ for $-\theta \le x \le +\theta$.

Assumptions 1 and 2 together imply the following stochastic linear difference equation model for changes in preference:

$$\mathrm{d}P(t+h) = [d - s \cdot P(t)] \cdot h + \varepsilon(t), \tag{2b}$$

where $\varepsilon(t)$ is a noise process with the following mean and variance:

$$E[\varepsilon(t) \mid P(t) = x]/h = E[\mathrm{d}P(t+h) \mid P(t) = x]/h - E[d - s \cdot P(t) \mid P(t) = x]$$
$$= (d - s \cdot x) - (d - s \cdot x) = 0,$$
$$V[\varepsilon(t) \mid P(t) = x]/h = V[\mathrm{d}P(t+h) - (d - s \cdot P(t)) \cdot h \mid P(t) = x]/h$$
$$= V[\mathrm{d}P(t+h) - (d - s \cdot x) \cdot h \mid P(t) = x]/h.$$

[3] More general models of growth are possible. For example, a quadratic growth model can be used:

$$\mu(x) = d + b \cdot x - c \cdot x^2.$$

The well-known logistic growth function, $\mu(x) = x \cdot [b - c \cdot x]$, is a special case of the quadratic growth model (set $d = 0$), and so is the linear growth model (set $b = -s$ and $c = 0$). Fortunately, it is relatively simple to develop the desired equations for choice probability and choice response time for any mean drift rate $\mu(x)$ which is a function of the position x. Thus, it is not necessary to commit to a particular growth function at this stage. Quadratic and linear growth models can be directly compared by chi-square difference tests. However, the linear growth model will receive special attention for two reasons: one is parsimony and the second is the close relation produced by this model and previous stochastic models of choice such as Thurstone's (1959) choice model and Ratcliff's (1978) diffusion model.

Note that the term $(d - s \cdot x) \cdot h$ is a constant (since x is a fixed number and not a random variable), and therefore this term has no effect on the variance in the above expression. This allows us to write

$$V[\varepsilon(t) \mid P(t) = x]/h = V[\mathrm{d}P(t+h) \mid P(t) = x]/h$$
$$= E[\mathrm{d}P(t)^2 \mid P(t) = x]/h - E[\mathrm{d}P(t) \mid P(t) = x]^2/h$$
$$= \varphi^2 - \mu_x^2 \cdot h.$$

At this point the stochastic linear difference equation model has six parameters: s = growth rate, d = mean difference in valence, ϕ^2 = diffusion rate, θ = the inhibitory bound, z = starting state, and h = the time unit. However, the time unit, h, is not estimated from the data. Rather, it is chosen a priori to be as close to zero as required to achieve an accurate approximation to the 'ideal' continuous time process. This leaves only the five free parameters to be estimated from the data, which is the same number of parameters as used in, for example, Ratcliff's (1978) diffusion model (including the drift rate variance parameter used in Ratcliff's model).

2.2. Choice probability

We begin by defining the set of preference states, denoted Ω, for the Markov chain. To determine the total number of states, the inhibitory bound, θ, is expressed[4] in terms of the step size, Δ, as $\theta = k \cdot \Delta$. In other words, the threshold for making a movement is k steps of size Δ away from the neutral state $P = 0$. The preference state, P, is an element of a state space, Ω, containing a total of $m = 2 \cdot k + 1$ states:

$$\begin{matrix} 1 & 2 & & k+1 & & m-1 & m \\ \Omega = \{-k\Delta, & -k\Delta + \Delta, & \ldots, & -\Delta, 0, +\Delta, & \ldots, & +k\Delta - \Delta, & +k\Delta\}. \end{matrix}$$

For convenience, the m states in Ω can be indexed s_i, $i = 1, \ldots, m$. The upper and lower bounds can be identified as states s_1 and s_m, where m is determined from θ as follows:

$$m = 2 \cdot (\theta/\Delta) + 1.$$

The initial starting position, $P(0) = z$, can be identified as state s_j, where j is determined from z as follows:

$$j = (z + \theta)/\Delta + 1.$$

The probability of choosing the alternative on the right, given that the starting state is s_j, equals the probability that the preference state reaches the upper bound, state s_m, before reaching the lower bound, state s_1. This probability is known to be

[4] This does, however, restrict the possible values of θ to integer multiples of Δ. The initial starting position is also restricted to integer multiples of Δ.

equal to the following expression (see Bhattacharya and Waymire, 1990, p. 234):

$$\Pr[\text{choose right}] = S(j-1)/S(m-1); \tag{3}$$

$$S(k) = \sum_i \varrho_i, \quad i = 1, 2, \ldots, k;$$

$$\varrho_i = r_1 \cdot r_2 \cdot r_3 \cdots r_i \quad (r_1 = 1);$$

$$r_i = p(-k\Delta + (i-1)\Delta)/q(-k\Delta + (i-1)\Delta), \quad i > 1;$$

and $p(x)$ and $q(x)$ are defined in (1a) and (1b).

Note that Pr[choose right] is an increasing function of the starting position, z (holding all other parameters constant). This is clear from the fact that increasing z only increases the numerator of (3).

In appendix A, we show that if the initial preference state is unbiased ($z=0$), and $\mu(x)>0$ for every x in Ω, then increasing the inhibitory bound by a step of magnitude Δ increases Pr[choose right] (holding all other parameters constant). For example, if we assume that $\mu(x)$ satisfies the conditions for the positive linear growth model, and that $d>0$ and $z=0$, then Pr[choose right] is an increasing function of the magnitude of the inhibitory bound, θ.

Also in appendix A we show that a decrease in the ratio $p(x)/q(x)$ for any given x in Ω causes an increase in Pr[choose right] (holding all other parameters constant). For example, if we assume that $\mu(x)$ satisfies the conditions for the positive linear growth model, then Pr[choose right] is an increasing function of the mean difference, d; and when $d>0$, Pr[choose right] is a decreasing function of diffusion rate, ϕ^2.

2.3. Distribution of choice response times

The transition probabilities, $p(x)$ and $q(x)$, can be collected together to form an $m \times m$ tridiagonal transition matrix, denoted T, with elements T_{ij}, $i=1,\ldots,m$, and $j=1,\ldots,m$, as follows. For the first and last row we set $T_{11}=T_{mm}=1$, otherwise zero. For all remaining rows, $T_{ii}=0$, and

$$T_{ij} = \begin{cases} p(-k\Delta + (i-1)\Delta) & \text{if } j-i=-1, \\ 1 - p(-k\Delta + (i-1)\Delta) - q(-k\Delta + (i-1)\Delta) & \text{if } j-i=0, \\ q(-k\Delta + (i-1)\Delta) & \text{if } j-i=+1. \end{cases}$$

The probability functions $p(x)$ and $q(x)$ were defined earlier in (1a) and (1b) in terms of the mean drift rate, $\mu(x)$, and the diffusion rate, ϕ^2.

The transition matrix T can be partitioned into five submatrices (see Appendix B): the top row, denoted A', contains the absorbing state for the left response; the bottom row, denoted B', contains the absorbing state for the right response; the rows 2 through $m-1$ in the first column, denoted R_1, contain the transition probabilities from each transient state to the left absorbing state; rows 2 through $(m-1)$

in the last column, denoted R_2, contain the transition probabilities from each transient state to the right absorbing state; and finally the $(m-2)\times(m-2)$ transient state matrix, denoted Q, formed by rows 2 through $(m-1)$ and columns 2 through $(m-1)$, contains the transition probabilities from one transient state to another transient state.

The initial starting distribution, represented by Z', is a $1 \times (m-2)$ row vector containing the initial probability distribution over the transient states. This distribution may be defined by a single parameter probability mass function such as the binomial distribution. However, if we assume a specific fixed starting position, $P(0)=z$, then all of the distribution in Z is concentrated on state s_j, where $j=(z+\theta)/\Delta+1$, corresponding to the initial preference state $P(0)=z$.

At this point we can make use of standard Markov chain theory (see Cox and Miller, 1965, ch. 3; Karlin and Taylor, 1981, ch. 10; Bhattacharya and Waymire, 1990, ch. 3) to derive the desired equations for choice probability and choice response time.

The probability of choosing alternative X ($X=1$ for left, $X=2$ for right) after deliberating for a time interval $t=(n+1)\cdot h$ equals

$$\Pr[\text{choose } X \text{ at time } t] = Z'Q^n R_X. \tag{4a}$$

This matrix computation can be greatly simplified by spectral analytic methods. It is well known (e.g. see Problem 5, p. 174, in Horn and Johnson, 1990) that the tridiagonal transition matrix Q is similar to a symmetric tridiagonal matrix (i.e. there exists a real valued matrix D such that $DQD^{-1}=S$, where S is a symmetrical tridiagonal matrix). This fact guarantees that Q has $m-2$ linearly independent eigenvectors and $m-2$ real eigenvalues (Searle, 1982, ch. 11; Horn and Johnson, 1990). Furthermore, by the Frobenius-Perron theorem (Cox and Miller, 1965, ch. 3; Karlin and Taylor, 1975, appendix; Bhattacharya and Waymire, 1990, ch. 3), all of the eigenvalues in Q are less than one in magnitude. Finally, efficient programs exist for finding the eigenvalues of tridiagonal symmetric matrices. Therefore, we factor Q as follows:

$$Q = P\Lambda P^{-1},$$

where P is the matrix of linear independent eigenvectors, and Λ is the diagonal matrix of real valued eigenvalues less than one in magnitude for the matrix Q. This allows us to write (4a) for $t=(n+1)\cdot h$ as

$$\begin{aligned} \Pr[\text{choose } X \text{ at time } t] &= (Z'P)\Lambda^n(P^{-1}R_X), \\ &= \sum_i v_i \cdot w_i \cdot (\lambda_i)^n, \quad i=1,\ldots,m-2, \end{aligned} \tag{4b}$$

where w_i is the ith coordinate of the row vector $Z'P$, v_i is the ith coordinate of the column vector $P^{-1}R_X$, and λ_i is the ith diagonal element of the diagonal matrix Λ. Equation (4b) is computationally efficient for large matrices.

The probability of choosing alternative X ($X=1$ for left, 2 for right) is obtained by summing over discrete time in (4a) to yield:

$$\Pr[\text{choose } X] = Z'\left[\sum_n Q^n\right]R_X, \quad n = 0, 1, 2, \ldots, \infty,$$
$$= Z'[I-Q]^{-1}R_X, \tag{5a}$$

where I is an identity matrix. If we define T as the random deliberation time, then the pth moment for the distribution of times to choose alternative X is obtained from (4a) by

$$E[T^p \mid \text{choose } X] = \frac{h^p \cdot Z'[\sum_n n^p \cdot Q^{n-1}]R_X}{\Pr[\text{choose } X]}, \quad n = 1, \ldots, \infty. \tag{5b}$$

In particular, the mean time equals

$$E[T \mid \text{choose } X] = \frac{h \cdot Z'[I-Q]^{-2}R_X}{\Pr[\text{choose } X]}.$$

The second moment equals

$$E[T^2 \mid \text{choose } X] = \frac{h^2 \cdot Z'[2 \cdot (I-Q)^{-3} - (I-Q)^{-2}]R_X}{\Pr[\text{choose } X]}.$$

(See Pike, 1966, for further discussion of the moment generating function for Markov chain models.)

In summary, choice probability and the entire distribution of choice response times can be computed quite easily on a personal computer using a matrix language such as, for example, GAUSS, MATLAB, MATHEMATICA, or SAS IML. For the linear growth model, only five parameters need to be estimated from the data (s, d, ϕ, z, k). These five parameters are used to define the state space, initial distribution, and transition probabilities of the Markov chain, which are then used to compute the distribution of choice response times. The time constant, h, is chosen to be as close to zero as needed to approximate a continuous time process. This value determines the dimension, m, of the transition matrix, which will be limited by the memory of the computer used to perform the calculations.

2.4. *Indifference responses*

In some experiments, subjects are allowed to express indifference rather than make a definite preference for one of the two alternative courses of action. Thus the decision-maker is asked to choose one from three options: prefer left, prefer right, or indifferent.

A simple model[5] for choice with indifference is to assume that an indifference response may occur with a small probability, denoted w, each time the preference state passes through the neutral state $P = 0$. Thus, it may be very unlikely that an indifference response will be made the first time that the preference state passes through the neutral point, but as the frequency of passing through the neutral state increases during deliberation, the likelihood of making an indifference response accumulates. If the preference state lingers around the neutral point for a long period of time, then the probability of an indifference response will be quite high.

More specifically, we attach an additional state, denoted I, to the original state space Ω for binary choice to produce a modified state space:

$$\Omega^* = \{-k\Delta, -k\Delta + \Delta, \ldots, -\Delta, 0, +\Delta, \ldots, +k\Delta - \Delta, +k\Delta\} \cup \{I\}.$$

A new $(m+1)\times(m+1)$ transition matrix, T^*, is formed from the original $m \times m$ transition matrix T as follows. The first m rows and the first m columns of the new matrix T^* are identical to the original matrix T, except for row $(k+1)$ of T, corresponding to the neutral preference state $P = 0$, which is multiplied by $(1 - w)$. The last row of T^* contains all zeros except for the last element corresponding to the new absorbing state, I, which is set to $T^*_{(m+1),(m+1)} = 1$. The last column of T^* contains all zeros except for the last row which is set to 1.0, and row $(k+1)$ corresponding to the neutral state which is set to $T^*_{(k+1),(m+1)} = w$.

The new transition matrix, T^*, can be partitioned into seven submatrices. The top row, labeled A', contains the absorbing state for choosing the left response; the second to last row, labeled B', contains the absorbing state for choosing the right response; the last row, labeled C', contains the absorbing state for the indifference response; rows 2 through $m-1$ in the first column, denoted R_1, contain the transition probabilities from the transient states to the absorbing state for the left response; rows 2 through $m-1$ in the second to the last column, denoted R_2, contain the transition probabilities from the transient states to the absorbing state for the right response; rows 2 through $m-1$ in the last column, denoted R_3, contain the transition probabilities from the transient states to the absorbing state for the indifference response; and, finally, the matrix formed by rows 2 through $m-1$ and columns 2 through $m-1$, denoted Q^*, contains the modified transition probabilities from one transient state to another.

The probability of choosing alternative X ($X = 1$ for left, 2 for right, 3 for indifferent) at time t for this indifference choice model is now given by (4a) except that the original transient state matrix, Q, used in (4a), is replaced by the modified transition matrix, Q^*. After making this same substitution in (5a) and (5b), one can also use these same equations to compute the choice probability and the moments for choosing each alternative in the choice with indifference problem. For example, the choice probabilities for the three responses are given by

[5] A more complex model would assume that an indifference response could occur from any state x in Ω with probability $w(x)$. The probabilities would be maximum for states in the neighborhood of the neutral state, and they would decrease as a function of distance from the neutral state. Fortunately, this more complex model remains mathematically tractable. The transient state matrix Q^* in (6) is obtained by multiplying each row of Q by $[1 - w(x)]$; and we change the row vector, R_3, in (6c) so that the element corresponding to state x has an exit probability equal to $w(x)$. Everything else in (6) remains the same. However, this generalization introduces additional new parameters, which we wish to avoid.

$$\Pr[\text{choose left}] = Z'[I - Q^*]^{-1}R_1, \tag{6a}$$

$$\Pr[\text{choose right}] = Z'[I - Q^*]^{-1}R_2, \tag{6b}$$

$$\Pr[\text{choose indiff}] = Z'[I - Q^*]^{-1}R_3. \tag{6c}$$

In summary, the choice response probabilities and choice response times for the choice with indifference problem can be computed from the linear growth model after estimating six parameters from the data $\{s, d, \phi, k, z, w\}$. The first five are identical to the parameters used in the binary choice model, and only one new parameter is needed for the indifference choice model.

2.5. Distribution of matched values

Suppose the decision-maker is presented with two alternatives, and he or she is asked to adjust the value of one dimension for one alternative until he or she is indifferent between the two alternatives. More specifically, suppose subjects are asked to state the cash difference between two alternatives by adjusting the money to be added or subtracted from the right alternative to make the subject indifferent between the left and right alternatives. The variable that is adjusted is denoted D, standing for the dial that is used to communicate the adjustments. The upper and lower bounds of the dial are $c \geq D \geq b$. These boundaries can be fixed by the experimenter.

We hypothesize that subjects perform a series of tests of dial values, and the dial value for the nth test is denoted $D(n)$. The dial remains fixed throughout a test period, during which a choice process is engaged. One of three events occurs during the choice process: a preference response favoring the left alternative may result, causing another test to be conducted using a new dial value that has been increased by an amount δ; a preference response for the right alternative may result, causing another test to be conducted using a new dial value that has been decreased by an amount δ; or an indifference response may result, causing the test process to stop, and the current dial value is reported as the matched value.

This matching process, $D(n)$, is a Markov chain process with transition probabilities

$$\Pr[D(n+1) = y - \delta \mid D(n) = y] = \Pr[\text{choose right} \mid D(n) = y] = u(y),$$

$$\Pr[D(n+1) = y + \delta \mid D(n) = y] = \Pr[\text{choose left} \mid D(n) = y] = v(y).$$

The probability of stopping the test process and reporting $D(n) = y$ as the final matched value equals $i(y) = 1 - u(y) - v(y)$.

For $b < y < c$, the probabilities $u(y)$ and $v(y)$ can be obtained from the indifference choice model as follows. Suppose the dial value is fixed at $D(n) = y$ during the nth test. This will affect the mean difference in valence, d, in the linear growth model (2a) because the dial value determines the amount of money added to or subtracted from the right alternative. Therefore the mean difference in valence is a function of the dial value, y. Then the probability of an increase, given that $D(n) = y$, $v(y)$, equals the right-hand side of (6a), with Q^* and R_1 functions of the dial value y. Similarly, the probability of a decrease given that $D(n) = y$, $u(y)$, equals the right-hand side of (6b) with Q^* and R_2 functions of the dial value y. At the boundaries, we assume that $u(c) = 1$ and $v(b) = 1$ (reflecting boundaries).

The state space for the dial values is defined in terms of the following two sets, with each set containing $J = (c - b)/\delta + 1$ states:

$$F = \{b^*, (b+\delta)^*, \ldots, (b + j\cdot\delta)^*, \ldots, (c-\delta)^*, c^*\},$$

$$G = \{b, (b+\delta), \ldots, (b + j\cdot\delta), \ldots, (c-\delta), c\}.$$

The first set, F, is a set of absorbing states representing all of the possible final matched values. In particular, state $(b + j\cdot\delta)^*$ refers to the case where the subject stops and reports a final matched value equal to $(b + j\cdot\delta)$. The second set, G, is a set of transient states representing all of the possible dial values that can occur during testing but prior to reporting the final matched value. In particular, $(b + j\cdot\delta)$ refers to the case where the subject tests the dial value $(b + j\cdot\delta)$ but does not report this value. The complete state space, containing $2\cdot J$ elements, is the union of these two sets, $\{F \cup G\}$.

The state transition probabilities $u(y)$, $v(y)$, and $i(y)$ can be collected together to form a canonical transition matrix M defined below:

$$M = \begin{bmatrix} V & W \\ 0 & I \end{bmatrix}.$$

The $J \times J$ transient state matrix V is defined as follows (see Appendix C):

$$V_{jk} = \begin{cases} u(b + (j-1)\cdot\delta) & \text{if } (k-j) = -1, \\ v(b + (j-1)\cdot\delta) & \text{if } (k-j) = +1, \\ 0 & \text{otherwise.} \end{cases}$$

The $J \times J$ submatrix W is a diagonal matrix:

$$W = \operatorname{diag}[0, \ldots, i(b + (j-1)\cdot\delta), \ldots, 0].$$

Finally, 0 is a $J \times J$ matrix of zeros, and I is a $J \times J$ identity matrix.

The probability distribution over the state space $\{F \cup G\}$ after n tests can be represented by a $1 \times 2\cdot J$ row vector denoted $Y(n)$. This probability vector can be partitioned into two parts, $Y_G(n)$, representing the distribution over the J transient states, and $Y_F(n)$, representing the distribution over the final absorbing states. $Y(0)$ represents the initial probability distribution, before testing begins ($n = 0$), and it is assumed that the decision-maker does not start in an absorbing state so that $Y_F(0) = 0$. The probability distribution after n tests is given by the product

$$Y(n) = Y(0)M^n.$$

$$= \left[Y_G(0)V^n \mid Y_G(0)\left(\sum_{j=0}^{n-1} V^j\right) W \right],$$

$$= [Y_G(0)V^n \mid Y_G(0)(I-V)^{-1}(I-V^n)W]. \tag{7a}$$

As n goes to infinity, we have $V^{\infty} = 0$, and (7a) converges to the following asymptotic distribution over the reported matched values:

$$Y_F = Y_G(0)[I - V]^{-1}W. \tag{7b}$$

In summary, for the matching task, three parameters can be fixed by the experimenter: b = the lower bound of the dial, c = the upper bound of the dial, and $D(0)$ = the initial dial setting. The distribution of final matched values for the matching task can be predicted from the linear growth model using only seven parameters that must be estimated from the data, $\{s, d, \phi, k, z, w, \delta\}$. The first five are identical to those used in the binary choice model, the sixth is the same as that used in the indifference choice model, and the last, δ, is the only new parameter, which represents the step size for the matching continuum in the matching task.

2.6. Model parameters

Consider an experiment which uses both a binary choice task (with an indifference response option included) and a selling price task for N choice stimuli. In a selling price task, subjects are asked to state the smallest amount of money for which they would be willing to sell a gamble to the experimenter. Suppose we estimate the $N \cdot (N-1)$ choice probabilities plus the $3 \cdot N \cdot (N-1)/2$ conditional mean choice response times from the choice task, and the N mean prices plus the N variances of the prices from the selling price task. This yields a total of $N(5N-1)/2$ data points.

For the linear growth model, four parameters may be estimated for each pair of stimuli: growth rate, s; mean difference in valence, d; diffusion rate, ϕ^2; and the initial preference state, $P(0) = z$. This produces a total of $2N(N-1)$ parameters for N stimuli. The number of steps from the neutral point to the threshold, k, is assumed to be constant across stimuli within a given task, but it may vary across tasks, producing two additional parameters. Finally, the indifference response probability, w, and the matching step size, δ, produce two more parameters. Altogether there are a total of $2N(N-1)+4$ parameters.

The difference between the number of data points and the number of parameters equals $N(N+3)/2-4$ degrees of freedom. Thus, the model is testable with only $N=2$ stimuli since $2(2+3)/2-4=1$. Additional tests of the model are possible by including higher moments such as the variance of the response times or by fitting the entire response time and matching distributions.

3. Continuous time and space diffusion model

Now we will assume that the preference state evolves continuously over time producing a continuous trajectory, $P(t)$. The state space is defined on the real interval $[-\theta, +\theta]$, and the time index set equals the non-negative reals.

The continuous-diffusion model is obtained from the discrete Markov chain model by letting the time unit, h, of the discrete time model approach zero in the limit (see Cox and Miller, 1965, pp. 213-215; Karlin and Taylor, 1981, pp. 168-169; Bhattacharya and Waymire, 1990, pp. 386-388). Although it is not explicitly shown, the mean drift rate, $\mu(x)$, and the diffusion rate, ϕ^2, are both functions of the finite time interval, h, for the discrete time model. Now we shall assume that the following limits exist:

$$\lim_{h \to 0} \mu(x) = \lim_{h \to 0} E[\mathrm{d}P(t) \mid P(t)=x]/h, \tag{8a}$$

$$\lim_{h \to 0} \phi^2 = \lim_{h \to 0} E[\mathrm{d}P(t)^2 \mid P(t)=x]/h. \tag{8b}$$

The first limit gives the infinitesimal drift rate, and second gives the infinitesimal diffusion rate. Hereafter, $\mu(x)$ and ϕ^2 will denote the inifinitesimal drift and diffusion rates, respectively.

When we let $h \to 0$, then equation (2b) becomes a stochastic differential equation called the Ornstein-Uhlenbeck (OU) process. It has infinitesimal drift rate $\mu(x) = d - s \cdot x$, and $\varepsilon(t)$ is a Wiener process with a mean of zero and variance ϕ^2. The close connection between the discrete Markov chain model and the continuous diffusion model is further developed in Appendix D.

The OU process was originally developed by physicists (Uhlenbeck and Ornstein, 1930), but more recently it has been applied to neurobiology (Ricciardi, 1977; Tuckwell, 1989) as well as reaction time research (Diederich, 1991; Smith, 1990, although Smith allowed for time dependent decision bounds) and risky decision-making (Busemeyer and Townsend, 1989). Here we will develop diffusion models of binary choice for two different types of decision tasks: an unspecified deliberation time task, and a specified deliberation time task.

3.1. Unspecified deliberation time model

In this subsection we return to the typical binary choice problem in which the deliberation time to reach a decision is a random variable determined by the time required for the magnitude of the preference state to exceed the inhibitory bound, θ, at which point a response is made. The right alternative is chosen as soon as $P(t) \geq +\theta$, and the left alternative is chosen as soon as $P(t) \leq -\theta$.

Define $u(t, z)$ as the probability density that the upper bound is reached for the first time at time t and that neither bound was reached before time t, given that the process started at state $P(0) = z$. As noted in Cox and Miller (1965, p. 230), $u(t, z)$

satisfies the backward equation (9):

$$(\partial u/\partial t) = (1/2)\cdot\phi^2\cdot(\partial^2 u/\partial z^2) + \mu(z)\cdot(\partial u/\partial z). \quad (9)$$

The solution to the backward equation (9) for the first passage time problem can be obtained by the separation of variables method (see Karlin and Taylor, 1981, p. 330). Suppose $u(t, z) = v(t)w(z)$. This solution is possible if we can find functions $v(t)$ and $w(t)$ that satisfy the pair of ordinary differential equations

$$\mathrm{d}v/\mathrm{d}t = -\lambda\cdot v,$$

$$\mu_z\cdot(\mathrm{d}w/\mathrm{d}z) + (1/2)\cdot\phi^2\cdot(\mathrm{d}^2 w/\mathrm{d}z^2) = -\lambda\cdot w,$$

where w satisfies the boundary conditions imposed by u, $w(+\theta) = w(-\theta) = 0$, for the finite boundary $\theta < \infty$.

This problem is a special case of the well known Sturm-Liouville problem (Boyce and DePrima, 1986, ch. 11). It is known that there exists a series of real valued eigenvalues, $\lambda_i < \lambda_{i+1}$, that increase toward positive infinity, and a series of real valued eigenfunctions, $w_i(z)$, which solve the above pair of ordinary differential equations so that the general solution can be expressed as a series solution:

$$u(t, z) = \sum_i v_i\cdot w_i(z)\cdot\exp(-\lambda_i\cdot t), \quad i = 1, 2, \ldots, \infty. \quad (10)$$

The coefficients v_i are chosen to satisfy the initial conditions for u at time $t = 0$, $u(0, z) = \delta(+\theta - z)$, where δ is the Dirac delta function.

The spectral analytic solution to the backward equation (10) resembles the spectral analytic solution for the discrete time model (4b). Both solutions are described by a weighted sum of eigenvalues. The main difficulty with (10) is the problem of finding the infinite series of eigenvalues and eigenfunctions that satisfy the initial conditions. This is why the discrete time Markov chain model is very practical because highly efficient programs are available which solve for the eigenvalues and eigenvectors of finite but large tridiagonal symmetric matrices. Furthermore, the solution from the discrete time model converges to the continuous time model, so that the continuous model can be approximated as closely as desired by the discrete model.

An alternative approach for obtaining the binary choice probabilities, and the moments of the deliberation time, is to use Laplace transform methods. The following is based on methods outlined by Cox and Miller (1965, ch. 5) and Karlin and Taylor (1981, ch. 15, pp. 191–204). First we need to provide some definitions. The joint probability that the stopping time is less than t and that the left alternative is chosen, given that the starting position is at $P(0) = z$, is defined here as

$$F(t, z) = \Pr[(T < t) \,\&\, \text{choose left} \mid P(0) = z].$$

The joint density for this distribution is defined as

$$(\partial/\partial t)F(t, z) = f(t, z).$$

Similarly,

$$G(t, z) = \Pr[(T < t) \,\&\, \text{choose right} \mid P(0) = z],$$

$$(\partial/\partial t)G(t, z) = g(t, z).$$

Finally,

$$H(t, z) = \Pr[T < t \mid P(0) = z] = F(t, z) + G(t, z),$$

$$h(t, z) = f(t, z) + g(t, z).$$

Hereafter we will focus on the derivation of the raw moments of the joint density, g. The derivation of the raw moments for the joint density f follows the same line of reasoning. The nth raw moment for g is defined as

$$m(n, z) = \int_0^\infty t^n\cdot g(t, z)\mathrm{d}t.$$

In particular, the raw moments that we desire are

$$m(0, z) = \Pr[\text{choose right} \mid P(0) = z],$$

$$m(n, z)/m(0, z) = E[T^n \mid \text{choose right}, P(0) = z], \quad n > 0.$$

The moment generating function for g is defined as

$$g^*(y, z) = \int_0^\infty \exp(-y\cdot t)\cdot g(t, z)\mathrm{d}t.$$

Recall the power series representation for $\exp(x)$:

$$\exp(x) = \sum x^n/n! \quad \text{for } n = 0, \ldots, \infty.$$

Substituting this into g^* yields the series representation for g^*:

$$g^*(y, z) = \sum (-y^n)/n!\cdot m(n, z), \quad n = 0, \ldots, \infty.$$

Thus, the raw moments of the stopping time distribution are obtained from g^* by $m(0, z) = g^*(0, z)$, and $m(n, z) = -(\partial^n/\partial y^n)g^*(y, z)$, evaluated at $y = 0$, for $n > 0$.

Now we will derive these raw moments from the backward equation of the OU process. Following Cox and Miller (1965, pp. 230–231), the Laplace transform of the backward equation implies that g^* satisfies the ordinary linear second-order differential equation shown below (for $-\theta < z < +\theta$):

$$((1/2)\cdot\phi^2)(\partial^2/\partial z^2)g^*(y, z) + (d - s\cdot z)\cdot(\partial/\partial z)g^*(y, z) = y\cdot g^*(y\cdot z). \quad (11)$$

Note that g^* also satisfies the following two boundary conditions:

$$g^*(y, -\theta) = 0, \text{ since } g(t, -\theta) = 0,$$

$$g^*(y, +\theta) = 1, \text{ since } m(0, +\theta) = 1 \text{ and } m(n, +\theta) = 0, \ n > 0.$$

Substituting the series representation for g^* into (11) yields

$$((1/2)\cdot\phi^2)(\partial^2/\partial z^2)\left[\sum(-y^n)/n!\cdot m(n,z)\right]$$
$$+(d-s\cdot z)\cdot(\partial/\partial z)\left[\sum(-y^n)/n!\cdot m(n,z)\right]$$
$$=y\cdot\left[\sum(-y^n)/n!\cdot m(n,z)\right]. \quad (12)$$

Equating coefficients of powers of y, yields a recursive series of equations:

$$((1/2)\cdot\phi^2)(\partial^2/\partial z^2)m(0,z)+(d-s\cdot z)\cdot(\partial/\partial z)m(0,z)=0. \quad (13a)$$

$$((1/2)\cdot\phi^2)(\partial^2/\partial z^2)m(n,z)+(d-s\cdot z)\cdot(\partial/\partial z)m(n,z)$$
$$=-n\cdot m(n-1,z),\quad n>0, \quad (13b)$$

with boundary conditions

$$m(0,-\theta)=0,\quad m(0,+\theta)=1,\quad m(n,-\theta)=0,\quad m(n,+\theta)=0.$$

Standard methods can be used to solve the differential equations (13a) and (13b). However, it is not necessary to do so here because the solutions to (13a) and (13b) can be obtained from Karlin and Taylor (1981, pp. 193–197) by setting their cost functional equal to our moment function, i.e. set $g(z)=n\cdot m(n-1,z)$ in their equation C. The solutions for $m(0,z)$ and $m(1,z)$ are:

$$m(0,z)=\Pr[\text{choose right}\mid P(0)=z]=S(z)/S(\theta),$$
$$m(1,z)/m(0,z)=E[T\mid\text{choose right},\ P(0)=z]$$
$$=2\cdot\{m(0,z)\cdot T_1(z)+[1-m(0,z)]\cdot T_2(z)\}/m(0,z),$$
$$S(x)=\int_{-\theta}^{x}\exp[(s\cdot y^2-2\cdot d\cdot y)/\phi^2]\mathrm{d}y,$$
$$f(x)=1/\{\phi^2\cdot\exp[(s\cdot x^2-2\cdot d\cdot x)/\phi^2]\},$$
$$T_1(z)=\int_{z}^{\theta}[S(\theta)-S(x)]\cdot f(x)\cdot m(0,x)\cdot\mathrm{d}x,$$
$$T_2(z)=\int_{-\theta}^{z}[S(x)-S(-\theta)]\cdot f(x)\cdot m(0,x)\cdot\mathrm{d}x.$$

3.2. Specified deliberation time model

Now we turn to a new type of decision task. Suppose the decision-maker i presented with a choice between two alternatives, but (s)he is told to make a commit ment at a specified or agreed-upon point in time, denoted t. Responses made prio to the commitment point are ignored, and only the response made at the agreed upon time point has any consequence. In this case, we assume that the magnitud

of the inhibitory process is set to an arbitrarily large[6] value ($\theta=\infty$) during the deliberation interval $[0,t]$, and then it is reduced to $\theta=0$ at the end of the interval immediately after time t. In this case the preference state continues to evolve in an unrestricted manner until the specified time point, t. Immediately after that point in time, the right alternative is chosen if $P(t)>0$ and the left is chosen if $P(t)<0$.

The choice probabilities for each specified time point t can be derived from the transition density of the unrestricted diffusion process. Define $u(t,z,y)$ as the probability density at the preference state y given that the process evolved for a duration t from an initial state $P(0)=z$. It is well known (see Cox and Miller, 1965, p. 215; Karlin and Taylor, 1981, p. 169; and Bhattacharya and Waymire, 1990, p. 388) that $u(t,z,y)$ satisfies the Kolmogorov backward equation (9). The transition density for the OU process is obtained by solving the backward equation (9) with the initial condition $u(0,z,y)=\delta(y-z)$ (where δ is the Dirac delta function), and with boundary conditions, $u(t,+\theta,y)=u(t,-\theta,y)=0$, for $t>0$. The solution of the unrestricted OU process model is obtained by setting the magnitude of the inhibitory bound to $\theta=\infty$. Under this condition the solution for the transition density is the normal density function with a mean $\eta(t)$ and variance $\upsilon^2(t)$ (Karlin and Taylor, 1981, p. 218):

$$u(t,z,y)=[2\pi\cdot\upsilon^2(t)]^{-(1/2)}\cdot\exp\{-(1/2)\cdot[(y-\eta(t))/\upsilon(t)]^2\}, \quad (14)$$
$$\eta(t)=\exp(-s\cdot t)\cdot z+[1-\exp(-s\cdot t)]\cdot(d/s),$$
$$\upsilon^2(t)=(\phi^2/2s)\cdot[1-\exp(-2\cdot s\cdot t)].$$

Finally, the probability of choosing the right alternative when given a fixed deliberation time t equals

$$\Pr[\text{choose right}\mid t]=F[\eta(t)/\upsilon(t)], \quad (15)$$

where F is the standard cumulative normal distribution function. The ratio

$$r(t)=\eta(t)/\upsilon(t)$$

is called the discriminability index. As the deliberation time approaches infinity, $r(t)$ approaches $r=(d/\phi)\cdot(\sqrt{2}/\sqrt{s})$. For $z=P(0)=0$,

$$r(t)=(d/\phi)\cdot\{(2/s)\cdot[1-\exp(-s\cdot t)]/[1+\exp(-s\cdot t)]\}^{(1/2)}. \quad (16)$$

4. Discussion

We conclude this paper by showing the relation of the decision field theory to several other models of decision-making. Busemeyer and Townsend (1989) give a longer and more detailed discussion of these relations. Here we wish to summarize

[6] The matrix methods developed for the discrete state, discrete time model are not useful for this specified deliberation time task because the matrices must be finite dimensional and we wish to use arbitrarily large boundaries for this task.

these relations more formally and concisely. For simplicity, we will focus on the equation giving the choice probability as a function of a specified deliberation time (15).

4.1. Expected utility model

Define X_i as a random variable representing the amount of money to be won or lost if gamble i is chosen. Define $u(x)$ as an increasing function of x, and $U_i = E[u(X_i)]$ is the expected value of the random variable $u(X_i)$, i.e. the expected utility of the gamble X_i. Suppose a decision-maker is presented with two gambles, one displayed on the left and the other displayed on the right. According to the classic expected utility hypothesis, the decision-maker always chooses the gamble with the larger expected utility. If $d^* = U_R - U_L$ is the difference between the expected utilities of the gambles on the right and left, then the gamble on the right is chosen if $d^* > 0$ and the gamble on the left is chosen if $d^* < 0$.

The expected utility model is an example of a deterministic and static model. Nevertheless, it can be obtained as a special case of decision field theory as follows. First, we set $\mu(x) = s \cdot (d^* - x)$, i.e. we use the proportional change growth model. Second, we set $d^* = U_R - U_L$, the difference in expected utility. Third, we allow the diffusion rate, ϕ^2, to approach zero. As ϕ^2 approaches zero, the function F in (15) approaches a step function—the probability of choosing the gamble on the right approaches zero or one depending on whether the sign of $\eta(t)$ is positive or negative, respectively. Letting the diffusion rate approach zero reduces all variability in the evolution of preference, which results in a deterministic dynamic model. Fourth, we allow the growth rate, s, to approach infinity. As s goes to infinity, $\eta(t)$ approaches a constant (independent of time) equal to $d^* = U_R - U_L$. Letting the growth rate approach infinity eliminates all of the dynamics to produce instantaneous choice. This highly constrained version of decision field theory satisfies all of the 'rational' axioms of expected utility theory (von Neumann and Morgenstern, 1947).

Of course, this 'ideal' model fails to describe the two most basic features of human decision making, namely that choice is probabilistic and that decisions take time. If we consider expected utility theory as a rational theory of decision-making, then decision field theory provides a way to parameterize and measure devisions from an 'ideal' decision-maker.

4.2. Additive difference model

According to Tversky's (1969) additive difference model, when given a choice between two n-dimensional choice alternatives, the probability of choosing the right alternative over the left is defined as

$$\Pr[\text{choose right}] = F\left[\sum \delta_i\right], \quad i = 1, \ldots, n,$$

where δ_i is the advantage or disadvantage for the right over the left alternative contributed by the ith dimension, and F is an increasing function that maps $\sum \delta_i$ into the closed interval [0, 1].

A special case of the additive difference model can be obtained from decision field theory as follows. First, we set $\mu(x) = s \cdot (d^* - x)$, with $d^* = \sum \delta_i$. Second, we allow s to approach infinity so that $\eta(t)$ approaches d^*. Finally, we set the diffusion rate, φ^2, equal to a fixed constant across all pairs of alternatives. Like the additive difference model, this version of decision field theory can produce violations of weak stochastic transitivity.

The additive difference model is more general than this special case of the specified deliberation time model because the function F for the additive difference model is only required to be monotonic. No process theory was postulated for the additive difference model, and consequently there is no reason for choosing one increasing function over another. For decision field theory, F must be set equal to the standard cumulative normal distribution function (see equation (15)). This is a logical consequence of assuming a stochastic linear differential equation model of choice (2). We would have to change our basic assumptions regarding the dynamics in order to derive an alternative choice probability function.

Considering the dynamic aspects, decision field theory is more general than the additive difference model, since the latter is static. Decision field theory is able to describe how the probability of choosing one alternative over another changes as a function of deliberation time. For example, suppose a customer is given a choice between a very familiar standard product (the status quo alternative) and a brand new product. At short deliberation times, the probability of choosing the status quo may exceed that for the brand new product, if the initial state of preference is biased toward the status quo. At longer deliberation times this probability may reverse if the new product is superior, so that the mean difference in valence favors the new product. Goldstein and Busemeyer (in press) provide a detailed discussion of an empirical example where choice probabilities changed from below to above (1/2) as a function of deliberation time.

4.3. Thurstone choice models

The general Thurstone choice model is obtained as a special case of (15) by setting (a) the mean difference in valence, d, equal to the difference between the means of the two alternatives, $d = U_R - U_L$, and (b) fixing the deliberation time, t, to be a constant. Like the Thurstone model, this version of the decision field model satisfies moderate stochastic transitivity (Halff, 1976).

The Thurstone model is static—the mean and variance are fixed across time; decision field theory is dynamic—the mean and variance both evolve as a function of deliberation time. This is a significant extension because it is possible to estimate the model parameters from an entire choice probability–deliberation time curve for a single choice pair (see, for example, Reed, 1973). This permits us to test the model