Multi-stage sequential sampling models: A framework for binary choice options Part 2

Adele Diederich

Jacobs University Bremen

Spring School SFB 1294
Dierhagen

March 18 – 22, 2019

Overview

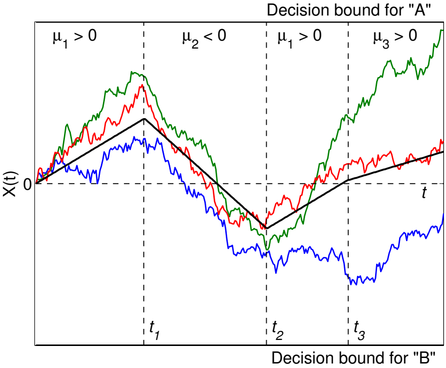
- Part 1
  - Motivation – 3 Examples
  - Basic assumptions of sequential sampling models (as used here)
  - Multi-stage sequential sampling models
- Part 2
  - Time and order schedules
  - Implementation
  - Predictions
  - Impact of attention time distribution
  - Impact of attribute order
- Part 3
  - Applications

Multiattribute choice alternatives: Basic assumptions

- For **each attribute** of the choice alternatives a **different** accumulation **process** takes place.
- Attention switches from attribute to another and attributes are processed **serially**

Multi-stage decision model

Example: Two choice alternatives with three attributes



μ_k drift rate for attribute k, direction of the process reflects strength of evidence

t_l attention switching

Schedules

- **Order schedule**: specific order in which attributes are considered
- **Time schedule**: specific times attention is switched from one attribute to another one
- Part of the model parameters

Times

- **Attention switching times**

$$T_0 = T_{start} = 0 < T_1 < T_2 < \ldots < T_L = T_{end} \leq \infty$$

- **Time horizon**: T_{end} maximum duration of the decision process
- $\Delta T_l = (T_{l-1}, T_l]$: the l-th **attention time interval**.

Schedules

- Finite number of attributes: $k = 1, \ldots, K$

Definition

A time and order schedule consists of a sequence $\{T_l\}_{l=1,\ldots,L}$ of attention switching times, and a sequence $\{k_l \in \{1, \ldots, K\}\}_{l=1,\ldots,L}$ of attribute indices which specifies that during the attention time interval ΔT_l the k_l-th attribute is considered. At attention switching time T_l, $l = 1, \ldots, L-1$, attention switches from attribute k_l to attribute k_{l+1}.

Process assumptions

- The sampling process $X(t)$ is specified by a sequence of attention switching times

$$T_0 = T_{start} = 0 < T_1 < T_2 < \ldots < T_{L-1} < T_L = T_{end} \leq +\infty$$

- The process $X(t)$ determined by such a schedule is a **piecewise** OUP or Wiener process.
- For $t \in [T_{l-1}, T_l]$ the process is determined by

$$dX(t) = (\delta_{k_l} - \gamma_{k_l} X(t))dt + \sigma dW(t)$$

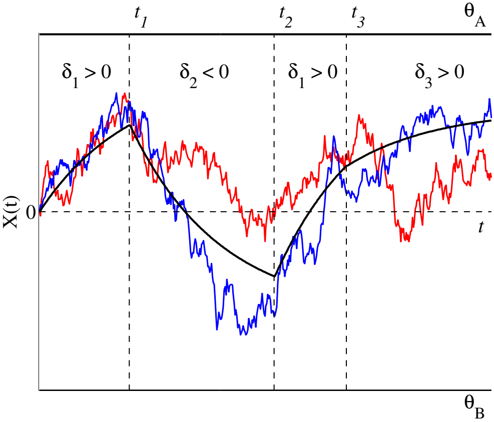
or

$$dX(t) = \mu_{k_l} - X(t)dt + \sigma dW(t)$$

Time and order schedule

- Time schedule: deterministic or random
- Order schedule: deterministic or random

Example: Deterministic time and order schedule of length $L = 4$



- Three different attributes ($K = 3$)
- Attribute order $(1, 2, 1, 3)$, i.e., $k_1 = 1$, $k_2 = 2$, $k_3 = 1$, $k_4 = 3$

Random time and order schedule

- **Time**: random variable T with a given distribution
- **Order**: stochastic $K \times K$ matrices $D^{(l)}$ such that $d^{(l)}_{k'k} \geq 0$ describes the probability with which attention switches from the k'-th attribute to the k-th attribute at switching time T_l, $l = 1, \ldots, L-1$

Order schedule – Example: Two attributes

- $d_{kk}^{(l)} = 0$ to avoid a no switching
- For two attributes ($K = 2$), this leads to $d_{11}^{(l)} = d_{22}^{(l)} = 0$ and $d_{12}^{(l)} = d_{21}^{(l)} = 1$
- The attribute sequence is either $(1, 2, 1, 2, \ldots)$ or $(2, 1, 2, 1, \ldots)$, depending on whether $k_1 = 1$ or $k_1 = 2$

Order schedule – Example: Three attributes

- For three attributes ($K = 3$) and $L = 3$, setting

$$D^{(1)} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}, \qquad D^{(2)} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 3/4 & 1/4 & 0 \end{bmatrix},$$

  results for $k_1 = 1$ in order sequences $(1, 2, 1)$, $(1, 3, 1)$, $(1, 3, 2)$ with probability $1/2$, $3/8$, $1/8$, respectively.
- Note, matrix $D^{(1)}$ models the situation when no preference or bias in switching from any given to any other attribute is assumed.

Order schedule – Example: Three attributes

- For three attributes ($K = 3$) and $L = 3$, setting

$$D^{(1)} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}, \qquad D^{(2)} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 3/4 & 1/4 & 0 \end{bmatrix},$$

results for $k_1 = 1$ in order sequence $(1, 2, 1)$ with probability $1/2$.

Order schedule – Example: Three attributes

- For three attributes ($K = 3$) and $L = 3$, setting

$$D^{(1)} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}, \qquad D^{(2)} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 3/4 & 1/4 & 0 \end{bmatrix},$$

results for $k_1 = 1$ in order sequence $(1, 3, 1)$ with probability $3/8$.

Order schedule – Example: Three attributes

- For three attributes ($K = 3$) and $L = 3$, setting

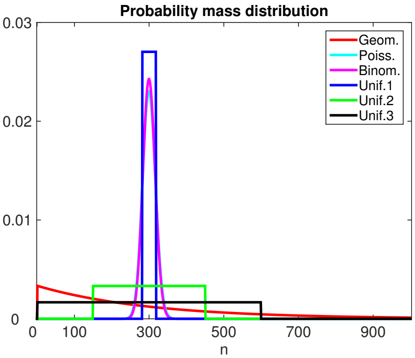
$$D^{(1)} = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}, \qquad D^{(2)} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 3/4 & 1/4 & 0 \end{bmatrix},$$

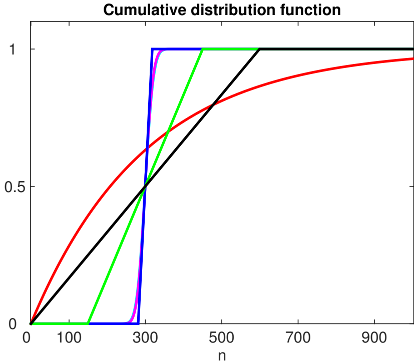
results for $k_1 = 1$ in order sequence $(1, 3, 2)$ with probability $1/8$.

Attention times

- Attention time, i.e., attention spent on an attribute, $\Delta T_L = [T_{l-1} - T_l]$
  - Deterministic
  - Geometric distribution $(Pr(\Delta T_L = n) = (1-r)^{n-1} r, n = 1, 2, \ldots)$
  - Poisson $(Pr(\Delta T_L = n) = e^{-\lambda} \frac{\lambda^n}{n!}, \qquad n = 0, 1, 2, \ldots,)$
  - Binomial $Pr(\Delta T_L = n) = \frac{N!}{n!(N-n)!} p^n (1-p)^{N-n}, \qquad n = 0, 1, \ldots, N,$
  - Uniform distribution $(Pr(\Delta T_L = n) = \frac{1}{2M+1}, n = N - M, \ldots, N + M)$
- Example
  - $E(\Delta T_L) = 300$
  - Unif 1: $M = \sqrt{N} \approx 18$
  - Unif 2: $M = N/2 = 150$
  - Unif 3: $M = N - 1 = 299$

Attention time distributions ΔT_L





Implementation

- Stochastic process approximation by discrete time, finite state space Markov chain
- Markov chain – matrix form

Deriving the Wiener process

- Brownian bridge; Kolmogorov-Smirnov statistics
- Feynman-Kac formula
- Stochastic Fourier analysis
- ...
- Scaled random walk (RW)

Scaled random walk

- The continuous state space $S = [\theta_A, \theta_A]$ of the Wiener process (or OUP) is replace by a finite state space

$$S = \{-m_B\Delta, \ldots, -2\Delta, -\Delta, 0, +\Delta, +2\Delta, \ldots, +m_A\Delta\},$$

with $\theta_A \approx m_A\Delta$ and $\theta_B \approx -m_B\Delta$, written as

$$S = \{x_i := i\Delta : \ i \in \mathcal{I}\}, \qquad \mathcal{I} = \{-m_B, \ldots, m_A\} \subset \mathbb{Z},$$

and $m = m_A\Delta + m_B\Delta + 1$ states.

- Δ: step-size of change in evidence

Scaled random walk

- The time space (parameter set) of a RW is $\{0, 1, 2, 3, \ldots\}$.
- Given a time interval $[0, t]$: Divide this into subintervals of length τ $\rightarrow t/\tau$ such subintervals
- The process makes a step at times $\tau, 2\tau, 3\tau, \ldots$.

Scaled random walk

- Let X_n denote the amount of information accumulated up to time unit $n \approx t/\tau$:

$$X_n = \sum_{i=1}^{n} \xi_i$$

- Assume that the distribution of the random variable ξ is i.i.d with

$$Pr[\xi_i = +\Delta] = Pr[\xi_i = -\Delta] = 0.5.$$

- Set $X_0 = 0$
- Set $\Delta = \sqrt{\tau}$

Standard Wiener process

- If $\tau \to 0$, then $X_{t/\tau}$ converges in distribution to a random variable W_t called *standard Wiener process* and has continuous time set and continuous state space.
- $W_t \sim \mathcal{N}(0, t)$

Wiener process with drift

- The random variable W_t is multiplied by a constant, σ, $\sigma > 0$, and a linear function of time, μt, is added, where μ, can be positive, negative or zero, and giving a particular initial value $X_0 = z$ results in

$$X_t = z + \mu t + \sigma W_t.$$

- $X_t \sim \mathcal{N}(z + \mu t, \sigma^2 t)$

OUP and birth-death chain

- Similar procedures work for the OUP approximation

MC matrix approach

- Wiener

  Probabilities to move up or down by Δ in small time interval τ, $\Delta = \sigma\sqrt{\tau}$

$$p^{(k)}_{i,i-1} = \frac{1}{2}\left(1 - \mu_k \frac{\sqrt{\tau}}{\sigma}\right), \qquad p^{(k)}_{i,i+1} = \frac{1}{2}\left(1 + \mu_k \frac{\sqrt{\tau}}{\sigma}\right),$$

- OUP

  Probabilities to move up or down by Δ in small time interval τ, $\Delta = \sigma\sqrt{\alpha\tau}$

$$p_{j,i} = \begin{cases} \frac{1}{2\alpha}\left(1 + (\delta_k - \gamma_k x_j)\frac{\sqrt{\tau}}{\sigma}\right), & j = i+1, \\ \frac{1}{2\alpha}\left(1 - (\delta_k - \gamma_k x_j)\frac{\sqrt{\tau}}{\sigma}\right), & j = i-1, \\ 1 - p_{i,i+1} - p_{i,i-1}, & j = i, \\ 0, & |j-i| > 1. \end{cases}$$

- kth attribute

MC matrix approach

$$P_k = \left[\begin{array}{c|c} I & 0 \\ \hline R_k & Q_k \end{array}\right]$$

$$= \left|\begin{array}{cc|ccccc} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ \hline p_{21}^{(k)} & 0 & p_{22}^{(k)} & p_{23} & \cdots & 0 & 0 \\ 0 & 0 & p_{32}^{(k)} & p_{33}^{(k)} & \cdots & 0 & 0 \\ 0 & 0 & 0 & p_{43}^{(k)} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & p_{m-3,m-2}^{(k)} & 0 \\ 0 & 0 & 0 & 0 & \cdots & p_{m-2,m-2}^{(k)} & p_{m-2,m-1}^{(k)} \\ 0 & p_{m-1,m}^{(k)} & 0 & 0 & \cdots & p_{m-1,m-2}^{(k)} & p_{m-1,m-1}^{(k)} \end{array}\right|$$

Submatrix and States

- **I**: Identity matrix, corresponds to the absorbing states associated with the two decision thresholds
- $\mathbf{Q}_k$: contains the transient probabilities, corresponding to the updating evidence process
- $\mathbf{R}_k$: contains the one-step transition probabilities from the transient to the absorbing states
- The first column vector of the matrix R_k (denoted by $R_{B,k}$) contains the transient probabilities for reaching alternative B, whereas the second $R_{A,k}$ contains the ones for alternative A

Initial distribution

- **Z**: starting position of the process over the transient states, S^*, before the decision process begins.
- Fixed state i_0, – the *initial state*, $Z_j = 1$ if $j = i_0$, $Z_j = 0$ if $j \neq i_0$.
  - Biases: $X_0 > 0$ or $X_0 < 0$
  - Unbiased: $X_0 = 0$
- Randomly located in in the state space according to a probability distribution **Z** on S^*, – *initial distribution*, $0 \leq z_i \leq 1$ and $\sum_i z_i = 1$
  - Biased: Probability mass on positions $j > (m-1)/2$ or on positions $j < (m-1)/2$
  - Unbiased: on position $j = (m-1)/2$

Distribution and moments for one attribute

$n = t/\tau$

- Probability distribution

$$Pr[T = n \cap \text{choose } A] = \mathbf{Z}'\mathbf{Q}^{n-1}\mathbf{R}_A, \qquad n = 1, 2, \ldots, \infty,$$

- Probability of choosing A

$$Pr(T_A < \infty) \approx p_A = \mathbf{Z}' \sum_{n=1}^{\infty} \mathbf{Q}^{n-1}\mathbf{R_A} = \mathbf{Z}'(\mathbf{I} - \mathbf{Q})^{-1}\mathbf{R}_A$$

- r-th moment for the distribution T to choose alternative A

$$E[T^r \mid A] \approx \frac{\tau}{Pr(A)} \mathbf{Z}' \sum_{n=1}^{\infty} n^r \mathbf{Q}^{n-1}\mathbf{R}_A$$

- Mean RT

$$E[T \mid A] \approx ET_A := \frac{\tau}{Pr(A)} \mathbf{Z}'(\mathbf{I} - \mathbf{Q})^{-2}\mathbf{R}_A$$

Multiple attributes – Deterministic time and order schedule

k attributes, L attention times

- **Probability** to choose A

$$p_A = Z' \sum_{i=1}^{n_1} Q_{k_1}^{i-1} R_{A,k_1} + Z' Q_{k_1}^{n_1} \sum_{i=n_1+1}^{n_2} Q_{k_2}^{i-(n_1+1)} R_{A,k_2} + \dots$$

$$\dots + Z' Q_{k_1}^{n_1} \dots Q_{k_{L-1}}^{n_{L-1}-n_{L-2}} \sum_{i=n_{L-1}+1}^{n_L} Q_{n_L}^{i-(n_{L-1}+1)} R_{A,k_L},$$

k attributes, L attention times

- **Mean RT** to choose A

$$ET_A = \frac{\tau}{p_A}\left[Z'\sum_{i=1}^{n_1} iQ_{k_1}^{i-1}R_{A,k_1} + Z'Q_{k_1}^{n_1}\sum_{i=n_1+1}^{n_2} iQ_{k_2}^{i-(n_1+1)}R_{A,k_2} + \ldots \right.$$
$$\left. \ldots + Z'Q_{k_1}^{n_1}\ldots Q_{k_{L-1}}^{n_{L-1}-n_{L-2}}\sum_{i=n_{L-1}+1}^{n_L} iQ_{n_L}^{i-(n_{L-1}+1)}R_{A,k_L}\right].$$

Multiple attributes – Random time schedule

- Probability mass distribution (pdf)

$$Pr(T_{at} = n) = p_{n,k}$$

- Cumulative distribution function (cdf)

$$Pr(T_{at} \leq n) = f_{n,k} := \sum_{i=0}^{n} p_{i,k}, \qquad n = 0, 1, \ldots$$

$L = 1$: Random time schedule – Probability

- $L = 1$, attribute k

$$
\begin{aligned}
p_{A,k} &= \sum_{n=1}^{\infty} p_{n,k} Z' \left(\sum_{i=1}^{n} Q_k^{i-1} \right) R_{A,k} \\
&= Z' \left[\sum_{i=1}^{\infty} \left(\sum_{n=i}^{\infty} p_{n,k} \right) Q_k^{i-1} \right] R_{A,k} \\
&= Z' \left[\sum_{i=0}^{\infty} (1 - f_{i,k}) Q_k^{i} \right] R_{A,k}
\end{aligned}
$$

Notation

- $p_{AB,k} := [p_{B,k},\ p_{A,k}]$

$$p_{AB,k} = Z' V_k, \qquad V_k := \left[\sum_{i=0}^{\infty} (1 - f_{i,k}) Q_k^i \right] R_k.$$

Example: Geometrically distributed attention time

- Geometric distribution

$$Pr(T_{at} = n) = (1 - r)^{n-1} r, \qquad n = 1, 2, \ldots,$$

$$\begin{aligned} V_k &= \sum_{i=0}^{\infty} \left(\sum_{j=i+1}^{\infty} r_k (1 - r_k)^{j-1} \right) Q_k^i R_k \\ &= \sum_{i=0}^{\infty} (1 - r_k)^i Q_k^i R_k \\ &= (I - (1 - r_k) Q_k)^{-1} R_k \end{aligned}$$

Example: Uniform distribution

- Uniformly distributed attention time

$$f_{i,k} = \frac{i - N + M + 1}{2M + 1} \text{ and } (1 - f_{i,k}) = \frac{M + N - i}{2M + 1}$$

(the surviver function).

$$V_k = \left(\sum_{i=0}^{N-M-1} Q_k^i + \sum_{i=N-M}^{N+M-1} \frac{N + M - i}{2M + 1} Q_k^i \right) R_k .$$

$L = 1$: Random time schedule – Expected time

- $L = 1$, attribute k

$$et_{A,k} = \sum_{n=1}^{\infty} p_{n,k} \left(\sum_{i=0}^{n-1} (i+1) Z' Q_k^i \right) R_{A,k}$$

$$= Z' \left[\sum_{i=0}^{\infty} \left(\sum_{n=i+1}^{\infty} p_{n,k} \right) (i+1) Q_k^i \right] R_{A,k}$$

$$= Z' \left[\sum_{i=0}^{\infty} (1 - f_{i,k})(i+1) Q_k^i \right] R_{A,k}.$$

$$et_{AB,k} := [et_{B,k},\ et_{A,k}] = Z' W_k$$

$$W_k := \left[\sum_{i=0}^{\infty} (1 - f_{i,k})(i+1) Q_k^i \right] R_k$$

$L = 2$: Distribuitons

- For each fixed $k_1 = k'$ and $n_1 = T'_{at}$ probabilities for reaching a decision after n_1 are given by

$$\left[Pr(T'_{at} < \frac{T_B}{\tau} < \infty), Pr(T'_{at} < \frac{T_A}{\tau} < \infty) \right]_{n_1 = T'_{at}, k_1 = k'}$$

$$\approx \sum_{k=1}^{K} d_{k'k} Z' Q_{k'}^{n_1} V_k = Z' Q_{k'}^{T'_{at}} (\sum_{k=1}^{K} d_{k'k} V_k)$$

$L = 2$: Probabilities

$$
\begin{aligned}
[p_B, p_A]_{k_1 = k'} &= Z' V_{k'} + \sum_{n \geq 0} p_{n,k'} Z' Q_{k'}^{n} \left(\sum_{k=1}^{K} d_{k'k} V_k \right) \\
&= Z' \left[V_{k'} + \left(\sum_{n \geq 0} p_{n,k'} Q_{k'}^{n} \right) \left(\sum_{k=1}^{K} d_{k'k} V_k \right) \right] \\
&= Z' \left[V_{k'} + B_{k'} \left(\sum_{k=1}^{K} d_{k'k} V_k \right) \right], \quad k' = 1, \dots, K,
\end{aligned}
$$

where

$$
B_k = \sum_{n \geq 0} p_{n,k} Q_k^{n}, \qquad k = 1, \dots, K,
$$

$L = 2$: Expected times

$$
\begin{aligned}
et_{AB}|_{k_1=k'} &= Z'W_{k'} + \sum_{n=0}^{\infty} p_{n,k} Z' Q_{k'}^{n} \left(\sum_{k=1}^{K} d_{k'k} (nV_k + W_k) \right) \\
&= Z' \left[W_{k'} + \left(\sum_{i=0}^{\infty} p_{i,k'} i Q_{k'}^{i} \right) \left(\sum_{k=1}^{K} d_{k'k} V_k \right) \right. \\
&\quad \left. + \left(\sum_{i=0}^{\infty} p_{i,k'} Q_{k'}^{i} \right) \left(\sum_{k=1}^{K} d_{k'k} W_k \right) \right] \\
&= Z' \left[W_{k'} + C_{k'} \left(\sum_{k=1}^{K} d_{k'k} V_k \right) + B_{k'} \left(\sum_{k=1}^{K} d_{k'k} W_k \right) \right],
\end{aligned}
$$

where

$$
C_k = \sum_{n \geq 0} p_{n,k} n Q_k^{n}, \qquad k = 1, \ldots, K.
$$

Notation summary

$$V_k := \left[\sum_{i=0}^{\infty}(1 - f_{i,k})Q_k^i\right] R_k$$

$$W_k := \left[\sum_{i=0}^{\infty}(1 - f_{i,k})(i+1)Q_k^i\right] R_k$$

$$B_k := \sum_{n \geq 0} p_{n,k} Q_k^n$$

$$C_k = \sum_{n \geq 0} p_{n,k} n Q_k^n$$

Order schedule of arbitrary lenght L

For arbitrary L, it is more convenient to write the resulting recursion in terms of *block-matrix-vector operations*.

- **Z** the $K \times 1$ array with each entry equal to the initial distribution Z for each of the K attributes(think of $\mathbf{Z}'$ as its transpose, a $1 \times K$ array with entries Z').
- **B** the $K \times K$ diagonal array with the B_k on the diagonal.
- **C** the $K \times K$ diagonal array with the C_k on the diagonal with the same attention time distributions corresponding to **B**.
- **I** the $K \times K$ diagonal array, with identity matrices I of the appropriate size on the diagonal.
- **V** the $K \times 1$ array with the V_k as entries according.
- **W** the $K \times 1$ array with the W_k as entries according.

Choice probabilities and mean choice response times for arbitrary L

$$\mathbf{p}_{AB} = \mathbf{Z}' \left((\mathbf{I} + \mathbf{B}D^{(1)}) \dots (\mathbf{I} + \mathbf{B}D^{(l-1)}) \right) \mathbf{V}$$

$$\mathbf{et}_{AB} = \mathbf{Z}' \left[\left((\mathbf{C}D^{(1)}) \dots (\mathbf{C}D^{(l-1)}) \right) \mathbf{V} + ((\mathbf{I} + \mathbf{B}D) \dots (\mathbf{I} + \mathbf{B}D))\, \mathbf{W} \right]$$

Notice

For the geometric distribution

- $B_k = r_k Q_k (I - (1 - r_k) Q_k)^{-1}$
- Closed form expressions also for: Poisson, binomial, uniform

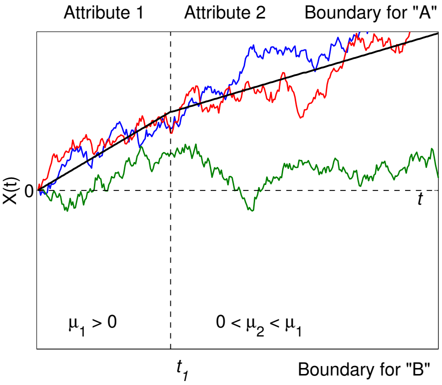
Predictions

- Parameters fixed
  - $\sigma = 1$; $\theta_A = -\theta_B = 10$
  - $\Delta = 1/4$, $\tau = 1/16 \rightarrow m = 81$ (matrix size)
  - $X(t) = 0$
- Expected value of attention time: $E(\Delta T_L) = 300$

Choice probabilities and choice response time

- Impact of attention time distributions
- Impact of attribute order

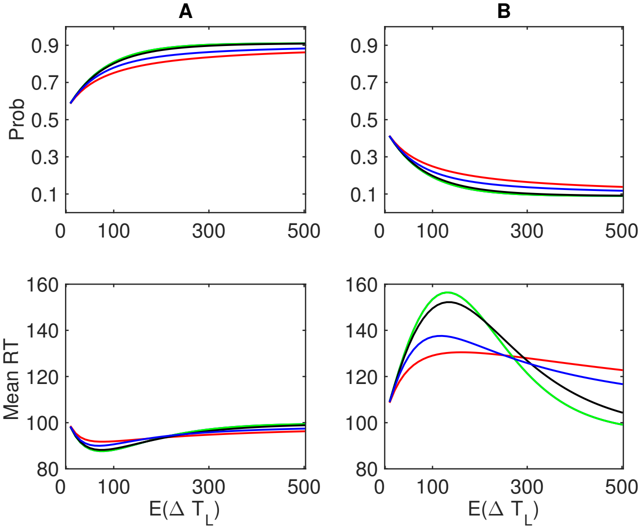
Case 1



- Both attributes favor choosing alternative A
- Attribute 1 with μ_1 is considered first
- Attribute 2 with μ_2 is considered second
- $\mu_1 > \mu_2$

Example 1: $\mu_1 = .1$; $\mu_2 = 0.01$; $k_1 = 1$; $k_2 = 2$; infinite

Det Geom Unif 1 Unif 2

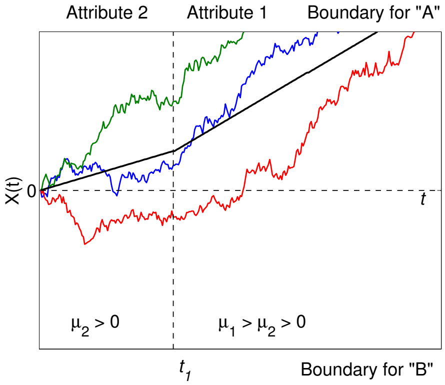


Example 1: Predicted choice pattern

$\mu_1 = .1 > \mu_2 = 0.01$; $k_1 = 1$; $k_2 = 2$;

The **more** frequently chosen alternative has **shorter** response times than the less frequently chosen alternative, regardless of the specific parameter values, and **regardless** of the underlying distribution for the attention time ΔT_L.

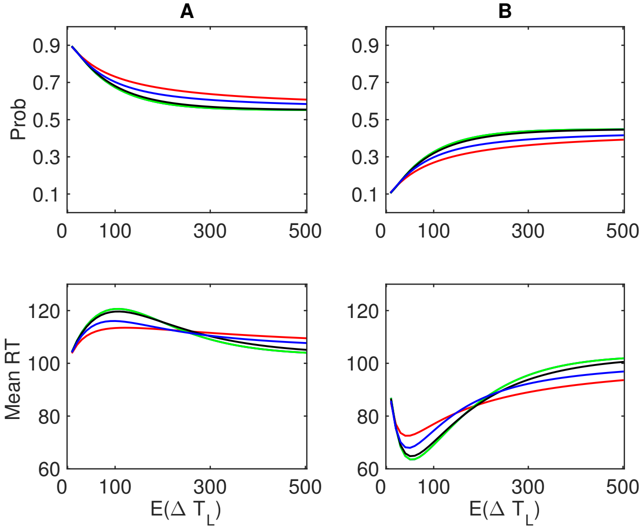
Case 2



- Both attributes favor choosing alternative A
- Attribute 2 with μ_2 is considered first
- Attribute 1 with μ_1 is considered second
- $\mu_1 > \mu_2$

Example 2: $\mu_1 = .1$; $\mu_2 = 0.01$; $k_1 = 2$; $k_2 = 1$; infinite

Det Geom Unif 1 Unif 2



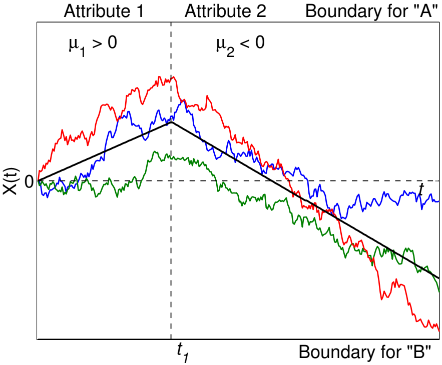
Example 2: Predicted choice pattern

$\mu_1 = .1 > \mu_2 = 0.01$; $k_1 = 2$; $k_2 = 1$;

The **more** frequently chosen alternative has **longer** response times than the less frequently chosen alternative, regardless of the specific parameter values, and **regardless** of the underlying distribution for the attention time T.

(**fast error**)

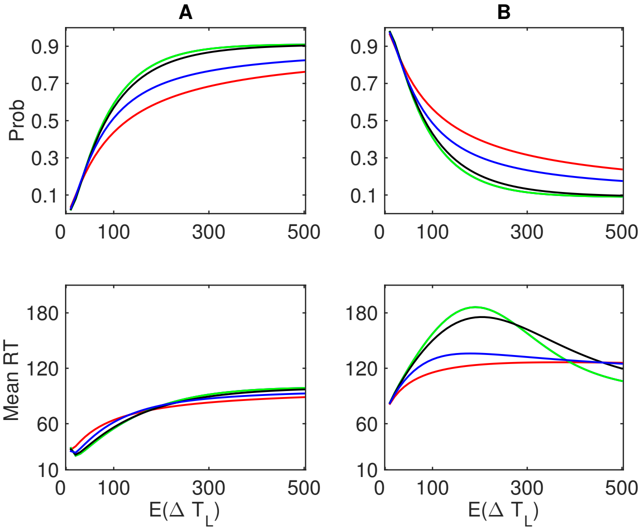
Case 3



- Attribute 1 with $\mu_1 > 0$ favors choosing alternative A
- Attribute 2 with $\mu_2 < 0$ favors choosing alternative B

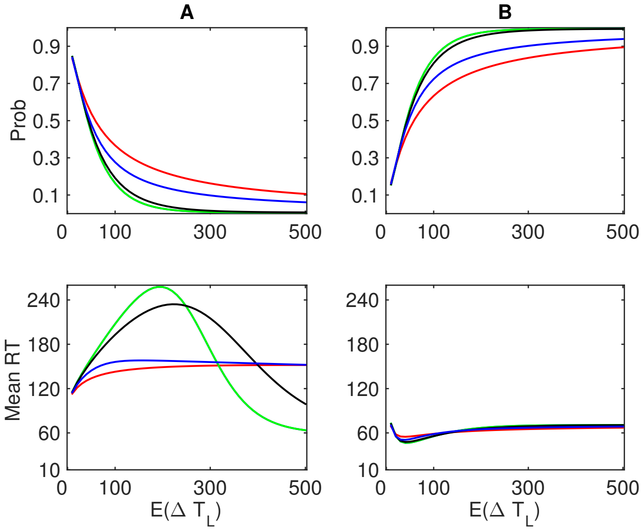
Example 3: $\mu_1 = .1$; $\mu_2 = -0.2$; $k_1 = 1$; $k_2 = 2$; infinite

Det Geom Unif 1 Unif 2



Example 4: $\mu_1 = .1$; $\mu_2 = -0.2$; $k_1 = 2$; $k_2 = 1$; infinite

Det Geom Unif 1 Unif 2



Examples 3 and 4: Predicted choice pattern

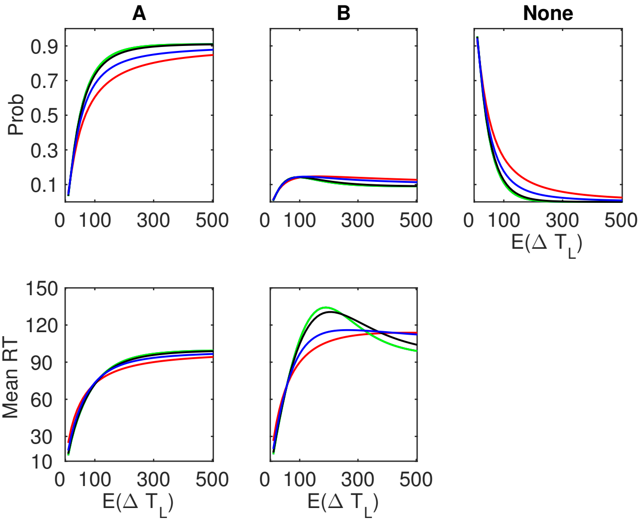
The model predicts preference reversals, regardless of the specific parameter values, and regardless of the underlying distribution for the attention time ΔT_L.
(and a rich choice response time/probability pattern)

Finite time horizon

- Examples 5, 6, and 7: similar to Examples 1, 2, and 3 but with finite time horizon

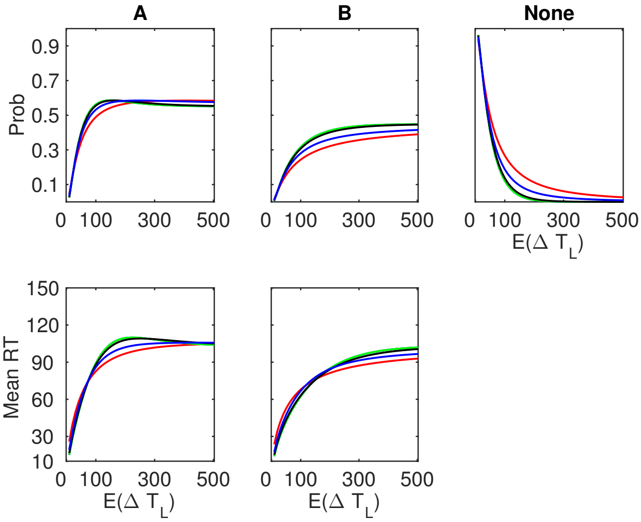
Example 5: $\mu_1 = .1$; $\mu_2 = 0.01$; $k_1 = 1$; $k_2 = 2$; finite

Det Geom Unif 1 Unif 2



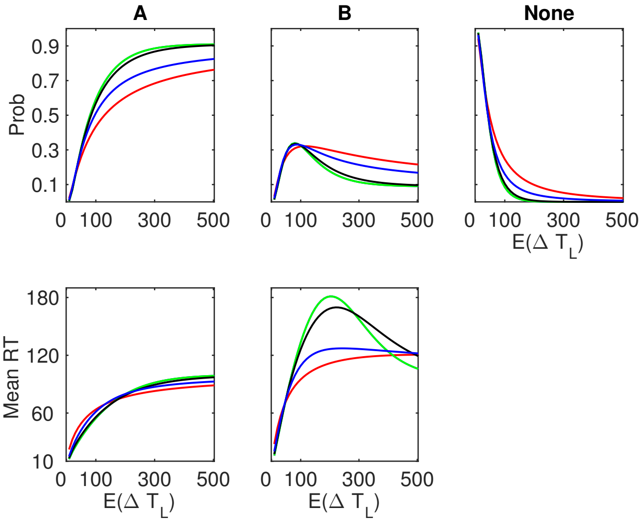
Example 6: $\mu_1 = .1$; $\mu_2 = 0.01$; $k_1 = 2$; $k_2 = 1$; finite

Det Geom Unif 1 Unif 2



Example 7: $\mu_1 = .1$; $\mu_2 = -.2$; $k_1 = 1$; $k_2 = 2$; finite

Det Geom Unif 1 Unif 2



Conclusion

- If $\mu_1 > \mu_2 > 0$ the model **always** predicts shorter response times for the more frequently chosen alternative, regardless of the assumed underlying attention time distribution.

Conclusion

- If $\mu_1 > \mu_2 > 0$ the model **always** predicts shorter response times for the more frequently chosen alternative, regardless of the assumed underlying attention time distribution.
- If $0 < \mu_1 < \mu_2$ the model **always** predicts faster responses for the less frequently chosen alternative, regardless of the assumed underlying attention time distribution.

Conclusion

- If $\mu_1 > \mu_2 > 0$ the model **always** predicts shorter response times for the more frequently chosen alternative, regardless of the assumed underlying attention time distribution.
- If $0 < \mu_1 < \mu_2$ the model **always** predicts faster responses for the less frequently chosen alternative, regardless of the assumed underlying attention time distribution.
- If $\mu_1 < 0 < \mu_2$ the model predicts preference reversals, regardless of the assumed underlying attention time distribution.

Conclusion

- With finite decision horizon, the model predicts a probability $p_0 > 0$ of not deciding, regardless of the assumed underlying attention time distribution.

Conclusion

- With finite decision horizon, the model predicts a probability $p_0 > 0$ of not deciding, regardless of the assumed underlying attention time distribution.
- The specific attention time distribution may be related to the experimental paradigm.
  - E.g., tracking eye movements: the sequence of attribute consideration and the switching times are directly observable $\longrightarrow$ deterministic or a uniform distribution with a small variance
  - E.g., all attributes are shown simultaneously (complex objects) and attention may shift at any moment in time $\longrightarrow$ a geometric distribution or a uniform distribution with a large variance

References

- Diederich, A. & Mallahi-Karai, K. (2018). Stochastic Methods for Modeling Decision-Making. In: Batchelder, W., Colonius, H., and E. N. Dzhafarov (Eds). New Handbook of Mathematical Psychology Vol. II Modeling and Measurement. Cambridge University Press.
- Diederich, A. & Oswald, P. (2014). Sequential sampling model for multiattribute choice alternatives with random attention time and processing order. *Frontiers in Human Neuroscience*, doi: 10.3389/fnhum.2014.00697