Bayesian Decision Theory
and
Sequential Sampling

Angela Yu

July 19, 2007

Introduction

Speed **vs.** **Accuracy**

- Faster responses save time, but more errors
- What is optimal policy?
- What computations are involved?
- Are humans/animals optimal?
- How is it neurally implemented?

Outline

- Review of Bayesian Decision Theory: examples
- Temporal cost: problem & solution
- Ex (1): 2-alternative forced choice
  - optimal policy (SPRT)
  - behavior & neurobiology
- Ex (2): 2AFC with finite deadline
  - optimal policy (SPRT + decaying threshold)
  - behavior & neurobiology

Bayesian Decision Theory

Decision policy

A **policy** $\pi : x \to d$ is a mapping from input x into decision d, where x is a sampled from the distribution $p(x|s)$.

Loss function:

The ***loss function*** $l(d(x); s)$ depends on the decision $d(x)$ and the hidden variable s.

Expected loss:

The ***expected loss*** associated with a policy π is averaged over the prior distribution $p(s)$ and the likelihood $p(x|s)$:

$$L(\pi) =< l(d(x); s) >_{x,s}$$

The ***optimal*** policy minimizes this expected loss.

Examples

Ex. 1: loss function is mean square error

$$l(\hat{s}(x); s) = (\hat{s} - s)^2$$

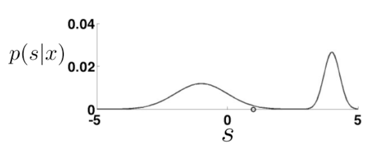
For each observation x, we need to minimize:

$$\langle l(\hat{s}(x); s) \rangle = \int (\hat{s} - s)^2 p(s|x) ds$$

Differentiating and setting to 0, we find loss is minimized if:

$$\hat{s}(x) = \int s p(s|x) ds = \langle s \rangle_{p(s|x)}$$

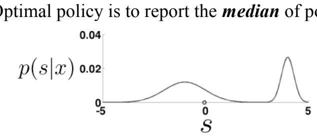
Optimal policy is to report the ***mean*** of posterior $p(s|x)$.



Examples

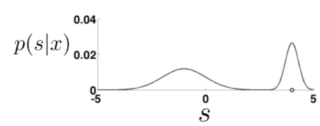
Ex. 2: loss function is absolute error: $|\hat{s} - s|$

Optimal policy is to report the ***median*** of posterior $p(s|x)$.

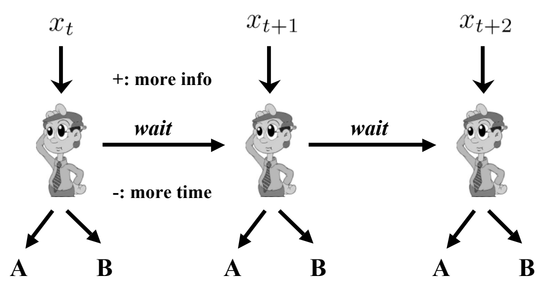


Ex. 3: loss function is 0-1 error: 0 if correct, 1 all other choices

Optimal policy is to report the ***mode*** of posterior $p(s|x)$.

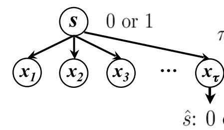


Time Matters → Repeated Decisions



An Example: Binary Hypothesis Testing

Decision Problem:



$$\pi(x_1, \ldots, x_t) \rightarrow \{0, 1, \text{cont}\}$$

Incorporate evidence iteratively (Bayes' Rule):

$$q_t \triangleq P(s{=}1|x_1, \ldots, x_t) = \frac{p(x_t|s{=}1)q_{t-1}}{p(x_t|x_1, \ldots, x_{t-1})}$$

Loss Function and Optimal Policy

Loss function penalizes both ***error*** and ***delay***:

$$l(\hat{s}, \tau; s) = \delta(\hat{s} - s) + c\tau$$

Expected loss:

$$L(\pi) = P_\pi(\hat{s} \neq s) + c\langle\tau\rangle_\pi$$

Optimal Policy (Bellman equation):

After observing inputs $x_1, \ldots, x_t$,

continue (observe x_{t+1}) only if the ***continuation cost < stopping cost***

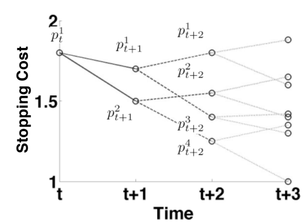
Stopping cost: $Q_s(q_t) = \min\{q_t, 1 - q_t\} + ct$

(q_t: $\hat{s}{=}0$; $1 - q_t$: $\hat{s}{=}1$)

Continuation cost: $Q_c(q_t) = \inf_{\tau > t}\langle l(\hat{s}, \tau; s)|q_t\rangle_{s,x}$

An Intractable Solution in Practice

$$Q_c(q_t) = \inf_{\tau > t}\langle l(\hat{s}, \tau; s)|q_t\rangle_{s,x}$$



Too hard!

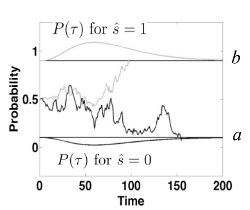
Useful Properties of Q_s and Q_c?

Wald & Wolfowitz (1948):

Continuation region is an ***interval*** (a, b) on the unit interval:

At time t, continue if $q_t < a$ and $q_t > b$,

stop and report $\hat{s}$=1if q_t>b, report $\hat{s}$=0 if q_t<a.



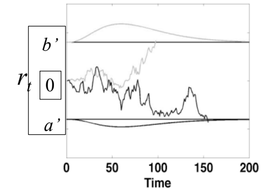
Sequential Probability Ratio Test

Log posterior ratio $r_t \triangleq \log \frac{q_t}{1-q_t}$ $\quad q_t = \frac{p(x_t|s{=}1)q_{t-1}}{Z_t}$

undergoes a random walk:

$$r_t = \log \frac{p(x_t|s{=}1)}{p(x_t|s{=}0)} + r_{t-1} \qquad r_0 = \log \frac{P(s{=}1)}{P(s{=}0)}$$

r_t is monotonically related to q_t, so we have (a',b'), a'>0, b'<0.



In continuous-time, a drift-diffusion process w/ absorbing boundaries.

Generalize Loss Function

We assumed loss is ***linear*** in error and delay:

$$L(\pi) = P_\pi(\hat{s} \neq s) + c\langle\tau\rangle_\pi$$

What if it's ***non-linear***, e.g. maximize reward rate:

$$\frac{1-P(\hat{s}\neq s)}{\langle\tau\rangle}$$ (Bogacz et al, 2006)

Wald also proved a ***dual*** statement:

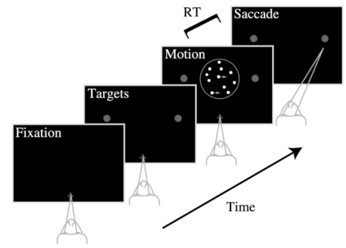
Amongst all decision policies satisfying the criterion $P(\hat{s} \neq s) \leq \alpha$

SPRT (with some thresholds) minimizes the expected sample size $<\tau>$.

This implies that the SPRT is optimal for ***all*** loss functions that increase with inaccuracy and delay (proof by contradiction).

Do People/Animals Behave Optimally?

A favorite task: Random dots coherent motion detection

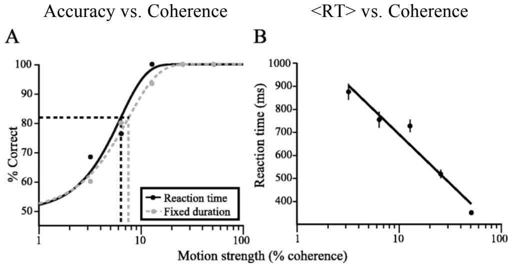


Properties: info/time slowed down, linear, and easily controlled

Do People/Animals Behave Optimally?

Accuracy vs. Coherence

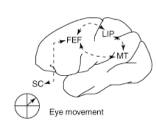
<RT> vs. Coherence



(Roitman & Shadlen, 2002)

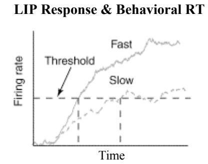
Neural Implementation

Saccade generation system

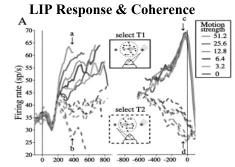


LIP - neural SPRT integrator? (Roitman & Shalden, 2002; Gold & Shadlen, 2004)

LIP Response & Behavioral RT



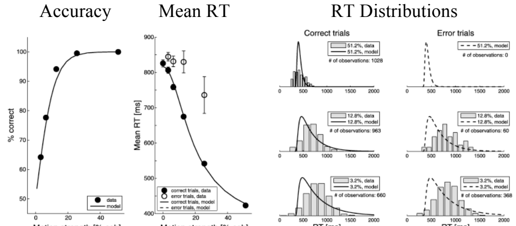
LIP Response & Coherence



Caveat: Model Fit Imperfect

(Data from Roitman & Shadlen, 2002; analysis from Ditterich, 2007)

Accuracy   Mean RT   RT Distributions

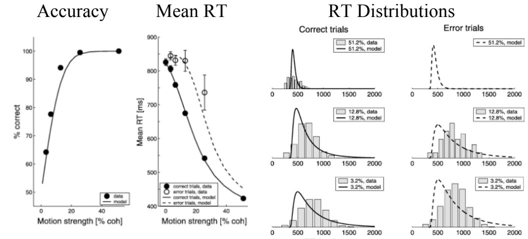


Fix 1: Variable Drift Rate

(Data from Roitman & Shadlen, 2002; analysis from Ditterich, 2007)

(idea from Ratcliff & Rouder, 1998)

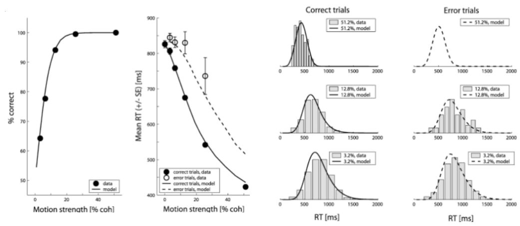
Accuracy   Mean RT   RT Distributions



Fix 2: Increasing Drift Rate

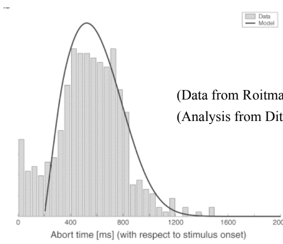
(Data from Roitman & Shadlen, 2002; analysis from Ditterich, 2007)

Accuracy Mean RT RT Distributions



A Principled Approach

Trials aborted w/o response or reward if monkey breaks fixation



(Data from Roitman & Shadlen, 2002)

(Analysis from Ditterich, 2007)

- More “urgency” over time as risk of aborting trial increases
- Increasing gain equivalent to decreasing threshold

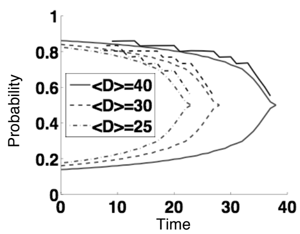
Imposing a Stochastic Deadline

(Frazier & Yu, 2007)

Loss function depends on error, delay, and whether deadline exceed

$$L(\pi) = P_\pi(\hat{s} \neq s)P(\tau < D) + c\langle\tau\rangle_\pi + dP(\tau \geq D)$$

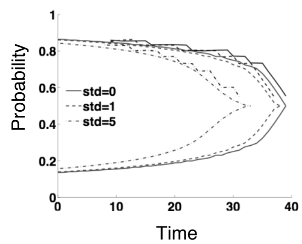
Optimal Policy is a sequence of ***monotonically decaying*** thresholds



Timing Uncertainty

(Frazier & Yu, 2007)

Timing uncertainty → more conservative policy (lower threhsolds)

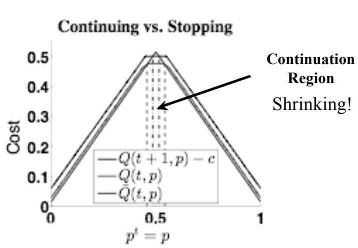


Theory applies to a large class of deadline distributions:

delta, gamma, exponential, normal

Some Intuitions for the Proof

(Frazier & Yu, 2007)



Summary

- Review of Bayesian Decision Theory: optimality depends on ***loss*** function
- Decision ***time*** complicates things: infinite repeated decisions
- Binary hypothesis testing, optimal policy is SPRT with fixed thresholds
- Behavioral & neural data ***suggestive***, but …
- … imperfectly fit. ***Variable/increasing*** drift rate both ad hoc accounts
- Impose deadline (with timing uncertainty) yields ***decaying*** thresholds
- Current/future work: generalize ***theory***, test model with ***experiments***