Decision Making Threshold

**Synonyms.**
Decision space boundary; absorbing boundary.

**Definition.**
A decision making threshold is the value of the decision making variable at which the decision is made, such that an action is selected or a commitment to one alternative is made, marking the end of accumulation of information.

**Description.**
A decision-making threshold determines when a decision-making process is completed. It represents a value of the decision variable, which in practice could be a linear combination of a set of neural firing rates, at which the accumulation of sensory evidence terminates and a response or action is chosen. In two-alternative forced choice tasks, two thresholds exist, one for each of the two alternatives.

In models of decision making, the threshold can be either static or dynamic during an individual trial. Mathematically, a threshold is an absorbing boundary of the diffusion process for the decision variable. The time for the decision variable to reach such an absorbing boundary is the decision-making time.

![Diagram](image)

**Figure 1.** Two static thresholds act as absorbing boundaries for the 1D Brownian motion of the decision variable, $X$. The boundary reached in any particular trial determines the choice (or percept in a perceptual categorization task) with the time for the decision variable to reach boundary comprising the dominant contribution to response time.

**Static Thresholds**
If the threshold is static, it determines where one is operating in the speed-accuracy tradeoff. Low thresholds produce speedier responses, whereas higher thresholds increase accuracy. In particular, for the drift diffusion model, which is based on a perfect integrator,
noise accumulates as the square root of time, whereas the signal accumulates linearly in time. Therefore, with large enough thresholds resulting in long enough decision-making time, accuracy of decisions can approach 100%. Specifically, if the decision variable has an initial value of zero and thresholds are set at $+a$ and $-b$, the two types of error rate, $\alpha$ (Probability of choosing + when the signal, $S$, is -) and $\beta$ (Probability of choosing - when the signal, $S$, is +) are given by,

$$
\alpha = \frac{e^{-\frac{2|S|b}{\sigma^2}} - 1}{e^{-\frac{2|S|b}{\sigma^2}} - e^{-\frac{2|S|a}{\sigma^2}}} \quad \text{and} \quad \beta = -\frac{1 - e^{-\frac{2|S|a}{\sigma^2}}}{e^{-\frac{2|S|a}{\sigma^2}} - e^{-\frac{2|S|a}{\sigma^2}}}.
$$

Notably, if thresholds are symmetric, with a starting point at the origin, then

$$
\alpha = \beta = \frac{1}{1 + e^{\frac{2|S|a}{\sigma^2}}},
$$

which decreases exponentially with increasing threshold, $a$.

Prior belief can be incorporated by setting the two thresholds to be different, reducing $a$ or $b$ respectively if a positive or negative response is preferable, or more likely, according to information acquired before stimulus onset.

Assuming a positive stimulus ($S > 0$) the distribution of response times for correct trials, $P(T^+)$, is given by

$$(1 - \beta)P(T^+) = \frac{\pi \sigma^2}{(a + b)^2} e^{\frac{Sb}{\sigma^2}} \sum_{k=1}^{\infty} k \sin \left( \frac{\pi ka}{a + b} \right) e^{-\frac{(S)^2 k^2 \sigma^2}{2(a+b)^2}} t$$

and the distribution of response times for incorrect trials, $P(T^-)$, is given by

$$\beta P(T^-) = \frac{\pi \sigma^2}{(a + b)^2} e^{\frac{Sa}{\sigma^2}} \sum_{k=1}^{\infty} k \sin \left( \frac{\pi kb}{a + b} \right) e^{-\frac{(S)^2 k^2 \sigma^2}{2(a+b)^2}} t.$$

This leads to a mean response time of (Bogacz et al., 2006):

$$
\langle T \rangle = (1 - \beta)\langle T^+ \rangle + \beta \langle T^- \rangle = \frac{a + b}{2S} \tanh \left[ \frac{S(a + b)}{2\sigma^2} \right] + \frac{(a + b) \left[ 1 - e^{\frac{S(a-b)}{\sigma^2}} \right]}{S \left[ e^{\frac{S(a+b)}{\sigma^2}} + e^{-\frac{S(a+b)}{\sigma^2}} \right]} + (b - a).
$$

In the symmetric condition, with $a = b$, then $P(T^+) = P(T^-)$ and the mean response times are given by:

$$
\langle T \rangle = \langle T^+ \rangle = \langle T^- \rangle = \frac{a}{S} \tanh \left( \frac{Sa}{\sigma^2} \right).
$$

If thresholds are static, they can be set at an optimal level, by making use of the above formulae, if the statistics of the task and the relative costs of different types of error are
known. In particular, if errors are costly, such as in lost time due to a long inter-trial interval, it pays to use thresholds far from the decision variable’s initial value.

**Figure 2.** In models of decision making with two variables, each variable represents a simple function (like the mean) of neural firing rates. **A)** If the thresholds correspond to a constant difference in firing rate, the model can map to a 1D model (Fig. 1) with static thresholds. **B)** If thresholds correspond to a constant absolute firing rate for each cell group, then a decision is made with smaller difference in rates as the sum of rates increases. This can correspond to a 1D model with thresholds, which collapse over a trial.

In models with two decision variables, a static threshold can depend either on each variable separately, independent of the value of the other variable, or it can depend on the difference between the two variables (see Figure 2). In terms of neural circuits, the former (Fig. 2A) corresponds to independent outputs from the competing groups of neurons, whereas the latter (Fig. 2B) corresponds to a push-pull arrangement, such that each group excites one response while inhibiting the other. In terms of equivalent single-variable models, it is notable that for a two-variable integrator model with static independent thresholds (Fig. 3), the difference in decision variables follows a single-variable integrator model with collapsing thresholds.

**Dynamic Thresholds**
In models of decision making with a single decision variable, dynamic thresholds can improve accuracy in tasks with fixed stimulus duration, and can optimize reward accumulation in other situations where the cost for acquiring information is not constant in time (Drugowitsch et al., 2012). If thresholds are initially high, they can ensure the decision process is not terminated too quickly and evidence has time to accumulate and outweigh the noise. However, if thresholds collapse to zero by the end of the stimulus duration, they
ensure a response is made and any evidence accumulated up until then contributes to the determination of that response.

When the statistics of the inputs are unknown, one can optimally adjust the thresholds over the course of the decision, according to whether the inputs are strongly informative or not. Perhaps surprisingly, when the inputs are more ambiguous in their evidence for one response or another, it is optimal for thresholds to collapse quickly, enforcing a rapid response that depends little on information in the stimulus, but more closely follows ones prior beliefs.
