Bachelor Thesis in Mathematics

Simulation of Mathematical Models of Decision Making

Department of Mathematics & Logistics
Jacobs University Bremen

Dania Sana

Bremen, May 17th, 2022

Supervised by Dr Keivan Mallahi-Karai

Statutory Declaration

Family Name, Given/First Name	Sana, Dania
Matriculation number	30002829
Kind of thesis sumbmitted	Bachelor thesis

English: Declaration of Authorship

I hereby declare that the thesis submitted was created and written solely by myself without any external support. Any sources, direct or indirect, are marked as such. I am aware of the fact that the contents of the thesis in digital form may be revised with regard to usage of unauthorized aid as well as whether the whole or parts of it may be identified as plagiarism. I do agree my work to be entered into a database for it to be compared with existing sources, where it will remain in order to enable further comparisons with future theses. This does not grant any rights of reproduction and usage, however.

This document was neither presented to any other examination board nor has it been published.

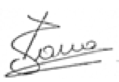
German: Erklärung der Autorenschaft (Urheberschaft)

Ich erkläre hiermit, dass die vorliegende Arbeit ohne fremde Hilfe ausschließlich von mir erstellt und geschrieben worden ist. Jedwede verwendeten Quellen, direkter oder indirekter Art, sind als solche kenntlich gemacht worden. Mir ist die Tatsache bewusst, dass der Inhalt der Thesis in digitaler Form geprüft werden kann im Hinblick darauf, ob es sich ganz oder in Teilen um ein Plagiat handelt. Ich bin damit einverstanden, dass meine Arbeit in einer Datenbank eingegeben werden kann, um mit bereits bestehenden Quellen verglichen zu werden und dort auch verbleibt, um mit zukünftigen Arbeiten verglichen werden zu können. Dies berechtigt jedoch nicht zur Verwendung oder Vervielfältigung.

Diese Arbeit wurde noch keiner anderen Prüfungsbehörde vorgelegt noch wurde sie bisher veröffentlicht.

Location, date: Bremen, 17 May 2022

Signature: Dania Sana



Abstract

Dynamical models of decision-making assume that during a decision-making process, information is being accumulated and once it passes a certain threshold, a decision is made. For example, the accumulation process of the standard binary dynamical model is given by a Wiener process with drift $\mathbf{X}(t)$. The choice is made and the process ends as soon as $|\mathbf{X}(t)| = \theta$, where θ denotes the positive threshold. The Cube model suggests that in a setting with d-alternatives, the accumulation process is given by the d-dimensional Wiener process with drift. This accumulation process is analyzed with respect to hitting the accept or reject boundaries of the alternatives, which determines the termination or continuation of the process. Disk model can only be applied to the setting with two alternatives and one fall-back option and its initial accumulation process is given by the 2-dimensional Wiener process with drift. Each of the two alternatives in this model, is characterized by reject and accept boundaries, that, when hit for the first time by the accumulation process, determine the termination of the process or continuation of it (as a circular Wiener process with drift). The least time needed for a choice to be made corresponds to the stopping time.

I simulate Cube model in the setting with three alternatives, two alternatives and the fall-back option as well as the Disk model. The main objective of my work was to detect the influence of the model parameters in the choice probabilities of the available options, stopping time (reaction time distribution) and episode durations. Lastly, I investigate, through the simulations, a new research direction in this field such as the Worst-best option ranking.

Contents

1 Introduction

The first part of this section introduces the Wiener process. The second part analyzes the relevance of this process in modelling decision making.

1.1 Theoretical Fundamentals

Wiener process (also known as the Brownian motion) is a continuous time random process which plays an important role in mathematical modelling. The term Brownian motion comes from the name of the botanist Robert Brown, who, in 1827, described the irregular motion of pollen grains suspended in the still water. The following rigorous definition of the Wiener process is due to Norbert Wiener, who studied it in the 1920s [1].

Definition 1 *The standard Wiener process* $\{W(t) : t \geq 0\}$ *is defined to be the unique process satisfying:*

1. $W(0) = 0$
2. *For* $0 \leq s \leq t$*, the increment* $W(t) - W(s)$ *is a Gaussian random variable with distribution* $\mathcal{N}(0,\ t - s)$
3. *The Gaussian random variables* $W(t_2) - W(t_1)$*,* $W(t_3) - W(t_2)$*, ...,* $W(t_r) - W(t_{r-1})$ *are independent for* $t_1 < t_2 < t_3 < ... < t_r$
4. $W(t)$ *is a continuous function of* t

The Wiener process is characterized by the properties described in the following theorem, which follows directly from Definition 1:

Theorem 1 *The Wiener process satisfies:*

1. *For* $s \geq 0$*, we have* $\mathbb{E}[W(s)] = 0$
2. *For* $s \geq 0$*, we have* $\mathbb{E}[W(s)^2] = s$
3. *For* $0 \leq s \leq t$*, we have* $\mathbb{E}[W(s)W(t)] = s$

In order to facilitate the reader understanding on this process, we will see a construction of it as the limit of the scaled symmetric simple random walk [2, p. 36-37]. First, let us define the simple random walk on the set of integers.

Definition 2 *A random walk is a stochastic sequence* S_n*, with* $S_0 = 0$*, defined by:*

$$S_n = \sum_{k=1}^{n} \varepsilon_k$$

where $\{\varepsilon_k\}$ *are independent and identically distributed random variables (i.i.d.). The random walk is simple if* $\varepsilon_k = \pm 1$*, with* $P(\varepsilon_k = 1) = p$ *and* $P(\varepsilon_k = -1) = 1 - p = q$*. Intuitively, consider a particle doing a random walk on the integer points of the real line which moves to one of its neighboring points in each unit time step; see Figure 1 [3].*

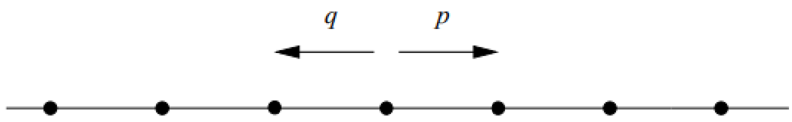


Figure 1: Simple random walk

To obtain the scaled symmetric simple random walk, start with a time interval $[0, t]$ and divide it into n subintervals of length $\tau = t/n$. Modify the simple random walk such that it makes a step at times $t = 0, \tau, 2\tau, 3\tau, ...$ In addition, instead of steps left and right of unit magnitude, suppose that the size of steps is either $+\Delta$ or $-\Delta$. We obtain the state space of this process as:

$$S = \{..., -2\Delta, -\Delta, 0, +\Delta, +2\Delta, ...\}.$$

Moreover, assume that the distribution of the independent identically distributed random variables ε_k is given by:

$$P[\varepsilon_k = +\Delta] = P[\varepsilon_k = -\Delta] = \frac{1}{2}.$$

For any $m \leq n$ with $s = m\tau \in [0, t]$, define:

$$X(s) = X(m\tau) := \varepsilon_1 + ... + \varepsilon_m.$$

Notice that:

$$\mathbb{E}[X(s)] = m \cdot \mathbb{E}[\varepsilon_1] = m(\frac{1}{2}\Delta - \frac{1}{2}\Delta) = 0 \text{ and } Var[X(s)] = m \cdot Var[\varepsilon_1] = m \cdot \Delta^2 = \frac{s}{\tau}\Delta^2$$

If we let τ and Δ be related by:

$$\Delta = \sqrt{\tau} \tag{1}$$

then, that for every $\tau > 0$ and every $s = m\tau$, where $0 \leq m \leq n$, we have:

$$\mathbb{E}[X(s)] = 0, \quad Var[X(s)] = mVar(\varepsilon_1) = s$$

Finally, if we let $\tau \to 0$, the random variable $X(s)$ converges in distribution to the *standard Wiener process* ($W(s)$). It is defined for all values $0 \leq s \leq t$ since the points $m\tau$ become increasingly denser in $[0, t]$ as $\tau \to 0$. Furthermore, since $W(s)$ is a Wiener process, $\mathbb{E}[W(s)] = 0$ and $Var[W(s)] = s$ which are also valid for $\mathbb{E}[X(s)]$ and $Var[X(s)]$, if equation (1) holds [2, p. 37-38].

The dynamical models of decision making that we will see in the following subsection, use a superimposition of the Wiener process with a deterministic linear process. In other words, for given constants μ and $\sigma > 0$, and an initial point z, we define the process

$$X(t) = z + \mu t + \sigma W(t) \tag{2}$$

to be the **Wiener process with drift**. It is a diffusion process which satisfies the stochastic equation given by:

$$dX(t) = \mu dt + \sigma dW(t)$$

The dynamical models of decision making that we will simulate, describe a setting with d alternatives, which implies the use of the *d-dimensional Wiener process* in the modelling.

Definition 3 *Let $W_1(t), W_2(t), ..., W_d(t)$ be standard independent Wiener processes. The vector-valued process*

$$\boldsymbol{W}(t) = (W_1(t), ..., W_d(t))^\top$$

is known as the d-dimensional Wiener process.

1.2 Dynamical Models of Decision Making

Decision making is the process of choosing an alternative from a set of alternatives. The process begins by gathering information, proceeds through likelihood estimation and deliberation and ends when a choice is made [4]. Dynamical models of decision making (known as sequential-sampling models) propose that after the onset of a stimulus, the decision maker accumulates information over time from the stimulus and/or its mental representation. The decision maker selects one of the available alternatives as soon as a decision criterion (specifying the amount of information required before a choice is made) is reached. Examples of dynamical models of decision making include race and diffusion models. Race models assume that evidence in favor of the available alternatives is accumulated in separate stores. The accumulation process stops whenever any of the stores first accumulates a predefined amount of evidence, and the corresponding alternative is chosen. The time at which this happens corresponds to the stopping time [5]. Diffusion models are typically applied to a binary choice task, so that evidence for one choice alternative is at the same time evidence against the other. A decision is made as soon as the process reaches one of two preset criteria [6]. It has been shown that a diffusion model (drift diffusion model) is applicable in economics since many reaction time phenomena in economics are consistent with the predictions from this model [7]. The following subsections present three dynamical models of decision making.

1.2.1 Standard binary model

Consider a decision setting with only two alternatives A_1 and A_2. The diffusion process $X(t)$ that satisfies the stochastic differential equation of the form $dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t)$ ($W(t)$ is the Wiener process as in Definition 1), represents the accumulation of information over time. This model assumes that the decision making process begins with an initial information available to the decision maker $X(0)$, which may either favor A_1 ($X(0) > 0$) or A_2 ($X(0) < 0$) or neither of them ($X(0) = 0$). The information is accumulated from one moment in time t to the next $t + h$ (h being small) by summing the current state with the new increment:

$$X(t + h) \approx X(t) + \mu(X(t), t)h + \sigma(X(t), t)(W(t + h) - W(t))$$

where $\mu(x, t)$ is called the drift rate and describes the deterministic rate of collecting information and $\sigma(x, t)$ is the diffusion rate which relates to the variance of the increments ($\mu(x, t)$ and $\sigma(x, t)$ are often assumed to be constant). This process continues until the magnitude of the cumulative evidence $X(t)$ exceeds a threshold criterion, θ. The process stops as soon as $X(t) = +\theta$ (alternative A_1 is chosen) or $X(t) = -\theta$ (option A_2 is chosen). $-\theta$ and $+\theta$ are also called **absorbing boundaries** of $X(t)$.

Important mathematical concepts of a decision making process are **stopping times** and **choice probabilities**. The stopping time refers to the minimal time t required for a choice to be made in a decision making process. For example, for this model, the stopping time can be expressed as:

$$\tau = \min\{t > 0 : \ X(t) = +\theta \text{ or } X(t) = -\theta\}$$

The random variable τ defines the **reaction time**. The choice probabilities for option A_1 and A_2 in t are given by:

$$p_{A_1} = P[X_\tau = +\theta] \ , \ \ p_{A_1} = P[X_\tau = -\theta] \tag{3}$$

It is worthy to note that the accumulation process of this model is such that any information collected in favor of A_1 is at the same time against the other alternative [6].

We will now present two models for the decision process with multiple alternatives introduced in [8] and [9] respectively.

1.2.2 Cube Model

Let the set $A = \{A_1, A_2, ..., A_d\}$ represent the possible alternatives in a decision making process. Cube model implies that the information accumulation at time t is given by a d-dimensional vector $\mathbf{X}$(t) whose i^{th} component indicates the amount of evidence collected in support of alternative A_i up to time t. This accumulation process is given by a Wiener process with drift in $\mathbb{R}^d$ (d-dimensional diffusion process) as:

$$\mathbf{X}(t) = \mathbf{x}_0 + \boldsymbol{\mu} t + \mathbf{\Sigma W}(t) \tag{4}$$

and it is the solution of the d-dimensional stochastic differential equation:

$$d\mathbf{X}(t) = \boldsymbol{\mu} dt + d\mathbf{\Sigma W}(t) \tag{5}$$

where $\mathbf{W}$(t), $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_d \end{pmatrix}$, and $\mathbf{\Sigma} = \begin{pmatrix} \sigma_{11} & \dots & \sigma_{1d} \\ \vdots & \ddots & \vdots \\ \sigma_{d1} & \dots & \sigma_{dd} \end{pmatrix}$ indicate the d-dimensional Wiener process as shown in Definition 3, the drift rate vector, and covariance matrix which controls the dependence structure of the d components of the noise vector, respectively. Moreover, $\boldsymbol{x}_0 = (x_1, ..., x_d)^\top$ is the initial tendency or inclination towards each alternative (when $x_i > 0$ the decision maker has an initially positive bias towards A_i). Each component of $\mathbf{X}(t)$ has the dynamics of a Wiener process with drift and we assume it is normalized so that $-1 \leq X_i(t) \leq 1$ for $1 \leq i \leq d$. Therefore, $\mathbf{X}(t)$ takes values in the d-cube:

$$\mathfrak{C}^d = \{(x_1, ..., x_d) : -1 \leq x_i \leq 1\}$$

The alternative A_j, $(1 \leq j \leq d)$ is characterized by the accept boundary:

$$\mathfrak{B}^+_{j,\,Cube} = \{(x_1, ..., x_d) \in \mathfrak{C}^d : x_j = 1\}$$

and the reject boundary:

$$\mathfrak{B}^{-}_{j,\,Cube} = \{(x_1, ..., x_d) \in \mathfrak{C}^d : x_j = -1\}.$$

The time until a choice has been made can be divided into **episodes**. The information accumulation process $\mathbf{X}(t)$ starts with $\mathbf{X}(0) = \mathbf{x}_0$ and continues until it hits one of the $2d$ boundaries for the first time. If, in this case, $\mathfrak{B}^{+}_{j,\,Cube}$ is hit, A_j is chosen and the decision making process ends in **one episode**. If instead, $\mathfrak{B}^{-}_{j,\,Cube}$ is hit, A_j is excluded from the set of available alternatives and the process continues as a $(d-1)$-dimensional diffusion process with the j^{th} component removed. The process is then analyzed with respect to hitting one of the $(2d-2)$ boundaries of the remaining available alternatives. Given that $d-1$ alternatives are excluded, the decision maker chooses the remaining alternative. This process development resembles the Tversky's elimination by aspects model according to which, alternatives are eliminated if they do not satisfy some basic criteria [10]. The **stopping time** of this process is the least time such that an alternative is chosen determining the termination of this process.

Let us illustrate this model in a setting with three alternatives ($d = 3$) and let the initial inclination towards each of the three alternatives be neutral, i.e., $\mathbf{x}_0 = (0,0,0)^\top$. The information accumulation process is described by:

$$\begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} t + \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} W_1(t) \\ W_2(t) \\ W_3(t) \end{pmatrix} \tag{6}$$

Assume that $\mathfrak{B}^{+}_{1,\,Cube}$ is hit first at a time t_1. The process ends in **one episode** and alternative A_1 is chosen. The stopping time of the process is t_1. If instead, $\mathfrak{B}^{-}_{1,\,Cube}$ is hit first, (at time t_2), then A_1 is excluded from being chosen and we are only left with A_2 and A_3. The accumulation process (starting from $t = 0$) will now be described by the 2-dimensional diffusion process:

$$\begin{pmatrix} X_2(t) \\ X_3(t) \end{pmatrix} = \begin{pmatrix} x_2(t_2) \\ x_3(t_2) \end{pmatrix} + \begin{pmatrix} \mu_2 \\ \mu_3 \end{pmatrix} t + \begin{pmatrix} \sigma_{22} & \sigma_{23} \\ \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} W_2(t) \\ W_3(t) \end{pmatrix} \tag{7}$$

This part of the process accounts for the **second episode** and starts where the **first episode** ended, therefore $x_2(t_2)$ and $x_3(t_2)$ denote the values of $X_2(t)$ and $X_3(t)$ at $t = t_2$ in (6). If the accept boundary of A_2 is hit first at time t_3 the process ends with the choice of A_2. The stopping time of the whole process (divided in two episodes) is $t_2 + t_3$. If instead, the reject boundary of A_2 is hit first at time t_4 the process finishes and option A_3 is chosen. The stopping time in this case is $t_2 + t_4$.

1.2.3 Cube model for two alternatives

In a setting with two alternatives $\{A_1, A_2\}$, the decision making process ends in one episode. If the accept boundary of, say, A_1 is hit for the first time, the process terminates with the choice of A_1. On the other hand, if the reject boundary of A_1 is hit initially, the process ends with the choice of A_2. Consider the following stopping times:

$$\tau_{A_1} = \min\{t : X(t) \in \mathfrak{B}^{+}_{1,\,Cube} \cup \mathfrak{B}^{-}_{2,\,Cube}\}, \quad \tau_{A_2} = \min\{t : X(t) \in \mathfrak{B}^{+}_{2,\,Cube} \cup \mathfrak{B}^{-}_{1,\,Cube}\}$$

The choice probability for A_1 correspond to the probability that the accumulation process hits $\mathfrak{B}^+_{1,\,Cube}$ or $\mathfrak{B}^-_{2,\,Cube}$ before hitting $\mathfrak{B}^+_{2,\,Cube}$ and $\mathfrak{B}^-_{1,\,Cube}$. One can mathematically express the choice probabilities for A_1 and A_2 by:

$$p_{A_1} = P[\tau_{A_1} < \tau_{A_2}], \quad p_{A_2} = P[\tau_{A_2} < \tau_{A_1}]$$

Process terminates as soon as one boundary of one of the alternatives is hit. In other words, if:

$$\overline{\tau_1} = \min\{t : |X_1(t)| = 1\}, \quad \overline{\tau_2} = \min\{t : |X_2(t)| = 1\}$$

the **reaction time**, which corresponds to the random variable determining the time it takes for a choice to be met, can be computed as:

$$\tau = \min(\overline{\tau_1}, \overline{\tau_2})$$

Lower and upper bounds for $\mathbb{E}[\tau]$ are estimated for a special case in the following theorem [8].

Theorem 2 *Consider the Cube model in a setting with two alternatives (d=2) whose accumulation process is given by* $\boldsymbol{X}(t) = (X_1(t), X_2(t))^\top$. *Its initial state corresponds to* $\boldsymbol{X}(0) = \boldsymbol{x}$. *With* τ *defining the reaction time, we have:*

$$\frac{1 - \|\boldsymbol{x}\|^2}{\|\boldsymbol{\Sigma}\|^2} \leq \mathbb{E}[\tau] \leq \frac{2 - \|\boldsymbol{x}\|^2}{\|\boldsymbol{\Sigma}\|^2}$$

where $\|\boldsymbol{x}\|^2 = \|(x_1, x_2)^\top\|^2 = x_1^2 + x_2^2$ *and* $\|\boldsymbol{\Sigma}\|^2 = \sum_{i,j=1}^{2} \sigma_{ij}^2$

1.2.4 Modification of the Cube Model: Fall-back option

Cube model can be modified by including the fall-back option in the set A of alternatives. This option is chosen if in every updated accumulation process, the boundary hit first is a reject boundary of one of the available alternatives. In other words, suppose that $d = 2$. Let $A = \{A_1, A_2, F\}$ denote the set of alternatives where F is the fall-back option. Suppose the initial inclination towards A_1 and A_2 is neutral. The accumulation process is described by:

$$\begin{pmatrix} X_1(t) \\ X_2(t) \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} t + \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix} \begin{pmatrix} W_1(t) \\ W_2(t) \end{pmatrix} \tag{8}$$

If $\mathfrak{B}^-_{1,\,Cube}$ is hit first at time t_{stop}, then A_1 is excluded from A. Suppose m is the value of X_2 at $t = t_{stop}$ in (8). The accumulation process (starting from $t = 0$) continues as a 1-dimensional Wiener process with drift:

$$X_2(t) = m + \mu_2 t + \sigma_{22} W_2(t)$$

If $\mathfrak{B}^-_{2,\,Cube}$ is hit first, this alternative is dropped from A and the fall-back option F is automatically chosen. Let us distinguish the following possible scenarios of this setting:

1. The accumulation process hits $\mathfrak{B}^+_{1,\,Cube}$ first; A_1 is chosen and the process ends.

2. The accumulation process hits $\mathfrak{B}^-_{2,\,Cube}$ first; A_2 is excluded from A. The process continues as 1-dimensional Wiener process with drift and it is considered with respect to hitting the boundaries of A_1:

- If $\mathfrak{B}^{+}_{1,\,Cube}$ is hit first, A_1 is chosen and the process ends.
- If $\mathfrak{B}^{-}_{1,\,Cube}$ is hit first, A_1 is excluded from A and the fall-back option is chosen.

3. The accumulation process hits $\mathfrak{B}^{+}_{2,\,Cube}$ first; A_2 is chosen and the process ends.

4. The accumulation process hits $\mathfrak{B}^{-}_{1,\,Cube}$ first; A_1 is excluded from A and the process continues as 1-dimensional Wiener process with drift and it is considered with respect to hitting the boundaries of A_2.

   - If $\mathfrak{B}^{+}_{2,\,Cube}$ is hit first, A_2 is chosen and the process ends.
   - If $\mathfrak{B}^{-}_{2,\,Cube}$ is hit first, A_2 is excluded and the fall-back option is chosen

1.2.5 Disk Model

Disk model is another dynamical model of decision making in a setting with two alternatives and the fall-back option. The accumulation process is initially described by the 2-dimensional Wiener process with drift given by (4). An important component of this model, which distinguishes it from the Cube model for two alternatives and fall-back option, is the **circular Wiener process**.

Definition 4 *Suppose $W(t)$ is a one-dimensional standard Wiener process and $\mu \in \mathbb{R}$, $\sigma > 0$. The process defined by*

$$\mathfrak{W}(t) = (cos(\theta_0 + \mu t + \sigma W(t)), sin(\theta_0 + \mu t + \sigma W(t))$$

is called the circular Wiener process with drift.

We consider $\{A_1, A_2, F\}$ to be the set of alternatives one chooses from, where F is the so-called fall-back option. Intuitively speaking, if one plans to buy a computer and needs to decide among two versions, there is the possibility of continuing working with the actual computer (fall-back option F). if these two computer versions (A_1, A_2) are not sufficiently attractive.

Similarly to the cube model, the alternatives A_1, A_2 are characterized by accept boundaries $(\mathfrak{B}^{+}_{1,\,Disk}, \mathfrak{B}^{+}_{2,\,Disk})$ and reject boundaries $(\mathfrak{B}^{-}_{1,\,Disk}, \mathfrak{B}^{-}_{2,\,Disk})$. We restrict the state space of the accumulation process to the unit disk $\mathfrak{D}$ by the inequality $x^2 + y^2 \leq 1$. Now, we fix an angle $0 < \alpha < \dfrac{\pi}{2}$ and divide the boundary $\mathfrak{D}$ into 4 circular walls. The accept and reject boundaries of alternative A_1, A_2 are obtained by the arcs:

$$\begin{aligned}
\mathfrak{B}^{+}_{1,\,Disk} &= \{\theta \in D : -\alpha \leq \theta \leq \alpha\},\\
\mathfrak{B}^{+}_{2,Disk} &= \{\theta \in D : \alpha \leq \theta \leq \pi - \alpha\},\\
\mathfrak{B}^{-}_{1,\,Disk} &= \{\theta \in D : \pi - \alpha \leq \theta \leq \pi + \alpha\},\\
\mathfrak{B}^{-}_{2,\,Disk} &= \{\theta \in D : \pi + \alpha \leq \theta \leq 2\pi - \alpha\}.
\end{aligned}$$

In the setting with two alternatives and one fall-option we encounter the following possible cases:

1. The 2-dimensional Wiener process with drift (4) hits $\mathfrak{B}^{+}_{1,\,Disk}$ for the first time; A_1 is chosen and the process ends.
2. The 2-dimensional Wiener process with drift hits $\mathfrak{B}^{-}_{2,\,Disk}$ first; A_2 is excluded from A. The process continues as a circular Wiener process with drift and it is considered with respect to hitting the boundaries of A_1:
   - If $\mathfrak{B}^{+}_{1,\,Disk}$ is hit for the first time, A_1 is chosen and the process ends.
   - If $\mathfrak{B}^{-}_{1,\,Disk}$ is hit first, A_1 is excluded from A and the fall-back option F is chosen.
3. The 2-dimensional Wiener process with drift hits $\mathfrak{B}^{+}_{2,\,Disk}$ first; A_2 is chosen and the process ends.
4. The 2-dimensional Wiener process with drift hits $\mathfrak{B}^{-}_{1,\,Disk}$ first; A_1 is excluded from A and the process continues as a circular Wiener process with drift and it is considered with respect to hitting the boundaries of A_2:
   - If $\mathfrak{B}^{+}_{2,\,Disk}$ is hit first, A_2 is chosen and the process ends.
   - If $\mathfrak{B}^{-}_{2,\,Disk}$ is hit first, A_2 is excluded and the fall back option is chosen

Similarly to the Cube model, the time until a choice has been made can be divided into **episodes**. For example, a decision is met in one single episode if either $\mathfrak{B}^{+}_{1,\,Disk}$ or $\mathfrak{B}^{+}_{2,\,Disk}$ is hit for the first time by the 2-dimensional diffusion process. However, if it hits for the first time, say $\mathfrak{B}^{-}_{1,\,Disk}$, one more episode is necessary to determine whether A_2 or fall-back option will be chosen. To summarize this model in the possible cases above, the underlying process in the first episode is described by the (2-dimensional) Wiener process with drift in (4). In scenarios (1) and (3) above, the decision making process is completed. In scenarios (2,4) the underlying process in the second episode is the circular Wiener process with drift. μ and σ are given, whereas θ_0 is chosen in such a way that the hitting point at the end of the first episode coincides with the initial point of the second episode.

The choice probabilities for each of the options in A have been found in [9] through the following proposition:

Proposition 1 *Consider the Wiener process issuing from the origin, that is, $\boldsymbol{X}(t) = \boldsymbol{W}(t)$. Let the components of the circular Wiener process be $\mu = 0$ and $\sigma = 1$. The choice probabilities for A_1, A_2, F are given by:*

$$P[A_{1,choose}] = \frac{1}{4} + \frac{\alpha}{2\pi}, \quad P[A_{2,choose}] = \frac{1}{4} + \frac{\pi/2 - \alpha}{2\pi}, \quad P[F] = \frac{1}{4}.$$

The proof of this proposition can be found in [9]. The mean of the first time the accumulation process $\mathbf{X}(t)$ (under specific parameters) hits the boundary of the unit disk, is estimated in the following proposition whose proof can be found in [9]:

Proposition 2 *Consider a Wiener process $\boldsymbol{W}(t)$ and set $\boldsymbol{X}(t) = \boldsymbol{\Sigma}\boldsymbol{W}(t) + \boldsymbol{x}$, where $\boldsymbol{x} \in \mathbb{R}^2$ is inside the unit circle D. Let τ denote the first time $\boldsymbol{X}(t)$ hits the boundary D. Then [9]:*

$$\mathbb{E}[\tau] = \frac{1 - ||x||^2}{||\Sigma||^2}.$$

The following section describes the necessary tools needed to simulate Cube mode with two alternatives and one fall-back option, Disk model, and Cube model with three alternatives.

2 Model Simulation

The main purpose of this thesis was to simulate the models described in the previous section and aim for numerical results concerning choice probabilities and reaction time. The fundamental process, needed to successfully implement the models, was the *d-dimensional Wiener process*. The tutorial [11] guides its implementation in `Python`. The realization of Cube model with two alternatives and fall-back option, Disk model, Cube model with three alternatives can be found in the **Appendix**.

2.1 Cube Model with two alternatives and fall-back option

The steps of this model are all defined in a function, that when called, displays the desired outcomes. If $\mathbf{X}(t)$ denotes the accumulation process, I consider the initial inclination of the decision maker towards the two alternatives to be neutral ($\mathbf{X}(0) = \mathbf{0}$) and components of the noise vector to be independent. We can then define the $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ to be:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix}$$

the corresponding function is:

>> `Cubemodel(t, N, n, d,` μ_1`,` μ_2`,` σ_{11}`,` σ_{22}`)`

where t corresponds to the end time in which the process runs, N defines the number of simulations, n is the number of time steps used, and d is the dimension of the the vector valued Wiener process.

2.2 Disk model

Similarly to the previous model, If $\mathbf{X}(t)$ is the accumulation process in the first episode, I let $\mathbf{X}(0) = \mathbf{0}$, and choose $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$, and α to be:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & 0 \\ 0 & \sigma_{22} \end{pmatrix}, \qquad \alpha = \frac{\pi}{4}$$

To obtain the simulation results of this model we call the function:

>> `Diskmodel(t, N, n, d,` μ_1`,` μ_2`,` σ_{11}`,` σ_{22}`,` μ`,` σ`)`

where t, N, n define the same variables as in the previous model, μ and σ are components of the circular Wiener process (accumulation process in the second episode).

2.3 Cube model with three alternatives

We consider the initial inclination of the decision maker towards each of the three alternatives to be neutral and take:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix}$$

We can call the following function to obtain the desired simulated results:

```
>> Cubemodel(t, N, n, d, μ1, μ2, μ3,  σ11, σ22, σ33)
```

where t, N, n, d are the same as defined in 2.1.

3 Simulation Results

This section contains the main results of my work in this thesis. I simulate the models and compare the proven conjectures in the first section with the simulated obtained results. I will in particular establish probabilities for events of interest, expected value of the time that determines the occurrence of such events, and the reaction time random variable of each simulated decision making model.

3.1 Reaction time and choice probabilities: Cube model with two alternatives and one fall-back option

Let us analyze the mean of the reaction time random variable and choice probabilities for each of the three options (A_1, A_2, F) of this model. Run the corresponding code found in **Appendix** which simulates the process 3000 times, considers the ending time and time step of the process to be 100 and 0.005 respectively. Option 1, option 2, and fall-back option obtained in the output of the code denote A_1, A_2 and F respectively. Take $\boldsymbol{\mu} = \mathbf{0}$, and $\boldsymbol{\Sigma} = \mathbf{I}$. We obtain:

```
>> Cubemodel(100, 3000, 20000, 2, 0, 0, 1, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.37475, 0.38425, and 0.241 respectively. The mean of reaction
   time is 1.1150782539126955.
```

We observe that the probability of choosing the fall-back option is near 0.25. This claim was proved in **Proposition 1** for the Disk model. However, up to a random change of time, these models are equivalent [9].

Consider another example where you increase the components of $\boldsymbol{\mu}$:

$$\boldsymbol{\mu} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

We obtain the results:

```
>> Cubemodel(100, 3000, 20000, 2, 2, 1, 1, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.68525, 0.31325, and 0.0015 respectively. The mean of reaction
   time is 0.41503575178758934.
```

The mean of the reaction time is much smaller than previously. The sign of the components of the drift vector determines the direction the process takes [2, pg. 41]. In other words, it impacts the tendency of the decision maker towards each of the alternatives. For example given:

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \qquad \mu_1 = -2 < 0, \quad \mu_2 = 1 > 0$$

we expect a much greater tendency of choosing option 2 rather than option 1 expressed in terms of choice probabilities:

```
>> Cubemodel(100, 3000, 20000, 2, -2, 1, 1, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.0095, 0.88525, and 0.10525 respectively. The mean of reaction
   time is 0.8186434321716085
```

This argument suggests that when the components of the drift vector are both negative, the process will move in the direction that avoids choosing option 1 or option 2, indicating the greatest choice probability for the fall-back option. To illustrate, take $\boldsymbol{\mu} = (-2, -1.2)^{\top}$, $\boldsymbol{\Sigma} = \mathbf{I}$ and run the code:

```
>> Cubemodel(100, 3000, 20000, 2, -2, -1.2, 1, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.011666666666666667, 0.07533333333333334, and 0.913 respectively.
   The mean of reaction time is 0.8435038418587595.
```

We see that that the fall-back option has the highest probability of being chosen. Let us see another example where $\boldsymbol{\mu} = (2, 0.5)^T$ and $\boldsymbol{\Sigma} = \mathbf{I}$. We obtain the choice probabilities and mean of reaction time as below:

```
>> Cubemodel(100, 3000, 20000, 2, 2, 0.5, 1, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.7796666666666666, 0.21366666666666667, and 0.006666666666666667
   respectively. The mean of the reaction time is 0.4498458256246146.
```

We observe that the probability of choosing option 1 is the greatest. Since $\mu_1 = 2 > \mu_2 = 0.5$, the process moves in the direction that tends to choose option 1. One can therefore conclude that the drift vector is strongly connected to the tendency towards the alternatives that the process exhibits.

Let us now analyze how the process behaves when changing the entries of $\boldsymbol{\Sigma}$. Choose:

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ 0.6 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 3 & 0 \\ 0 & 4 \end{pmatrix}.$$

We have the following output:

```
>> Cubemodel(100, 3000, 20000, 2, 1, 0.6, 3, 4)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.37066666666666664, 0.419, and 0.21033333333333334 respectively.
   The mean of the reaction time is 0.10923379502308447.
```

Consider another example where:

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ 0.6 \end{pmatrix}, \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}.$$

```
>> Cubemodel(100, 3000, 20000, 2, 1, 0.6, 1, 4)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.439, 0.5043333333333333, and 0.056666666666666664 respectively.
   The mean of the reaction time is 0.4492157941230394.
```

In both cases, option 2 has the highest probability of being chosen whereas the fall-back option hast the least one. Moreover, when I decrease σ_{11} from 3 to 1, we notice an increase of the mean of reaction time in this case. Change the input parameters to:

$$\boldsymbol{\mu} = \begin{pmatrix} 1 \\ 0.6 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}.$$

```
>> Cubemodel(100, 3000, 20000, 2, 1, 0.6, 3, 2)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.4803333333333333, 0.336, and 0.18366666666666667 respectively.
   The mean of the reaction time is 0.20559361301398402.
```

The results demonstrate another tendency from previously; option 1 has the highest choice probability and the mean of reaction time is the smallest in this case. Now, consider the components of the drift vector to be equal and find the results for the following cases:

$$\boldsymbol{\mu}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}; \qquad \boldsymbol{\mu}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}; \qquad \boldsymbol{\mu}_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{\Sigma}_3 = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

We obtain the following results for each case:

```
>> Cubemodel(100, 3000, 20000, 2, 1, 1, 1, 2)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.40766666666666667, 0.5496666666666666, and 0.042666666666666665
   respectively. The mean of the reaction time is 0.45249429138123576.

>> Cubemodel(100, 3000, 20000, 2, 1, 1, 2, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.5356666666666666, 0.422, and 0.042333333333333334 respectively.
   The mean of the reaction time is 0.45296931513242333.

>> Cubemodel(100, 3000, 20000, 2, 1, 1, 2, 3)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.36866666666666664, 0.4683333333333333, and 0.163 respectively.
   The mean of the reaction time is 0.20265013250662534.
```

We see that when the drift vector has the same components, increasing σ_{22} (σ_{11}) will imply a greater tendency of choosing option 2 (option 1). Moreover, in this case, we observe that increasing the components of the diagonal matrix $\boldsymbol{\Sigma}$ yields a smaller mean of reaction time.

3.2 Reaction time and choice probabilities: Disk model

Disk model describes the decision making setting with two alternatives and one fall-back option. As outlined in the previous section, [9] derives semi-closed formulas for the choice probabilities of each option for $\boldsymbol{\mu} = \mathbf{0}$, $\boldsymbol{\Sigma} = \mathbf{I}$, $\mu = 0$, and $\sigma = 1$. The validity of my implementation is tested by comparing the simulation-obtained choice probabilities with the theoretically-proved ones. Simulate the process 3000 times and let the ending time and time step be 100 and 0.005 respectively. Take as initial inputs: $\boldsymbol{\mu} = \mathbf{0}$, $\boldsymbol{\Sigma} = \mathbf{I}$, $\mu = 0$, and $\sigma = 1$. Execute the code pertaining to this model and obtain the simulation and theoretical results:

```
>> Diskmodel(100, 3000, 20000, 0, 0, 1, 1, 0, 1)[0]
>> The choice probabilities for option 1, option 2, and fall back option
   are 0.383, 0.37633333333333335, and 0.24066666666666667 respectively.
   The mean of reaction time is 0.8009283797523209. According to the
   paper [9], the choice probabilities for option 1, option 2, and
   fall-back option are 0.375, 0.375, and 0.25 respectively
```

Let us analyze the resulting choice probabilities and mean of the reaction time when changing the drift vector $\boldsymbol{\mu}$ and the components of $\boldsymbol{\Sigma}$. Consider $\boldsymbol{\Sigma} = \mathbf{I}$, $\mu = 0$, $\sigma = 1$ and increase the magnitude of $\boldsymbol{\mu}$ by taking $\boldsymbol{\mu} = (1, 2)^\top$.

```
>> Diskmodel(100, 3000, 20000, 1, 2, 1, 1, 0, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.31633333333333336, 0.663, and 0.020666666666666667 respectively.
   The mean of reaction time is 0.3756037801890094.
```

Similarly to the previous model, the mean of the reaction time has decreased because we increased the components of the drift vector (from $(0, 0)^\top$ to $(1, 2)^\top$). Furthermore, the probability for choosing option 2 is the highest compared to the choice probabilities for option 1 and fall-back option. This result is closely related to the magnitude of the second component of the drift vector which imposes a greater tendency for choosing option 2. Let us now increase the first component of the drift vector (take $\boldsymbol{\mu} = (2, 0)^\top$) and observe its impact on the choice probability for option 1:

```
>> Diskmodel(100, 3000, 20000, 2, 0, 1, 1, 0, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.7976666666666666, 0.15, and 0.052333333333333336 respectively.
   The mean of reaction time is 0.4593763021484407
```

Let us now take $\boldsymbol{\mu} = (-2, 0)^\top$. Based on our previous argument on the Cube model in the same setting, we expect a tendency of choosing option 2 over option 1:

```
>> Diskmodel(100, 3000, 20000, -2, 0, 1, 1, 0, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.06333333333333334, 0.49366666666666664, and 0.443 respectively.
   The mean of reaction time is 0.8650349184125874..
```

Let us simulate this model again and vary the components of $\boldsymbol{\Sigma}$. Consider three cases:

$$\boldsymbol{\mu}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ \boldsymbol{\Sigma}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}; \qquad \boldsymbol{\mu}_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ \boldsymbol{\Sigma}_2 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}; \qquad \boldsymbol{\mu}_3 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \ \boldsymbol{\Sigma}_3 = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

```
>> Diskmodel(100, 3000, 20000, 0, 0, 1, 2, 0, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.2966666666666667, 0.466, and 0.23733333333333334 respectively.
   The mean of reaction time is 0.5265213260663033.

>> Diskmodel(100, 3000, 20000, 0, 0, 2, 1, 0, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.4683333333333333, 0.293, and 0.23866666666666667 respectively.
   The mean of reaction time is 0.5202876810507191.

>> Diskmodel(100, 3000, 20000, 0, 0, 2, 3, 0, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.3333333333333333, 0.42433333333333334, and 0.24233333333333335
   respectively. The mean of reaction time is 0.3960931379902328.
```

Similarly to the Cube model (where we used $\boldsymbol{\mu} = (1,1)^\top$) we notice that when $\boldsymbol{\mu} = 0$, increasing the σ_{22} (σ_{11}) impacts the probability of choosing option 2 (option 1) by increasing it. Interestingly, the three above cases (in which we increased the components of $\boldsymbol{\Sigma}$), show a smaller mean of reaction time compared to the first example where $\boldsymbol{\mu} = \mathbf{0}$, $\boldsymbol{\Sigma} = \mathbf{I}$, $\mu = 0$, and $\sigma = 1$.

The second episode of this process depends on μ and σ which are defined in the circular Wiener process. This process is precisely the image of the one-dimensional Wiener process with drift under the map $W \to (cosW, sinW)$. Therefore, using this process is not different from using the Wiener process [9]. One should therefore expect, that when increasing μ, (and not changing other parameters), the mean of reaction time will decrease. Let us illustrate this with two examples:

$$\boldsymbol{\mu_1} = \boldsymbol{\mu_2} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\Sigma_1} = \boldsymbol{\Sigma_2} = \begin{pmatrix} 1 & 0 \\ 0 & 1.2 \end{pmatrix}, \quad \mu_1 = 0, \quad \mu_2 = 10, \quad \sigma_1 = \sigma_2 = 1.$$

```
>> Diskmodel(100, 3000, 20000, 0, 0, 1, 1.2, 0, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.3596666666666667, 0.393, and 0.24733333333333332 respectively.
   The mean of reaction time is 0.7351917595879796.

>> Diskmodel(100, 3000, 20000, 0, 0, 1, 1.2, 10, 1)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.5193333333333333, 0.29533333333333334, and 0.18533333333333332
   respectively. The mean of reaction time is 0.5016217477540544.
```

The results support our claim, however the change in the mean of reaction time is not very meaningful (compared to the choice of μ_1 and μ_2 that have a difference of 10 units).

Let us run two other examples, in which only σ changes:

$$\boldsymbol{\mu_1} = \boldsymbol{\mu_2} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{\Sigma_1} = \boldsymbol{\Sigma_2} = \begin{pmatrix} 1 & 0 \\ 0 & 1.2 \end{pmatrix}, \quad \mu_1 = \mu_2 = 0, \quad \sigma_1 = 1, \quad \sigma_2 = 5.$$

We already have the result for the first example. Let us run the second one:

```
>> Diskmodel(100, 3000, 20000, 0, 0, 1, 1.2, 0, 5)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.3383333333333333, 0.4053333333333333, and 0.25633333333333336
   respectively. The mean of reaction time is 0.4696801506742004.
```

Let us have another example with $\sigma = 10$:

```
>> Diskmodel(100, 3000, 20000, 0, 0, 1, 1.2, 0, 10)[0]
>> The choice probabilities for option 1, option 2, and fall-back option
   are 0.35133333333333333, 0.4076666666666667, and 0.241
   respectively. The mean of reaction time is 0.45451272563628176.
```

In these examples we observe that increasing σ (while keeping all other parameters as defined), yields a smaller mean of reaction time.

3.3 Reaction time and choice probabilities: Cube model with three alternatives

Simulate the process 3000 times and choose the ending time and time step of the process to be 500 and 0.01 respectively. Let us run the code to find the choice probabilities and mean of the reaction time for $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \mathbf{I}$.

```
>> Cubemodel(500, 3000, 50000, 3, 0, 0, 0, 1, 1, 1)[0]
>> The choice probabilities for option 1, option 2, and option 3 are
   0.319, 0.3473333333333333, and 0.33366666666666667 respectively.
   The mean of the reaction time is 0.737754755095102.
```

We observe that there is an equal probability for choosing each of the three options. This should not be a surprise, since the Cube model in this scenario is symmetric with respect to each of the alternatives.

Let us now take $\boldsymbol{\Sigma} = \mathbf{I}$ and $\boldsymbol{\mu} = (1, 2, 1)^T$. Reasoning as before, we claim that the mean of the reaction time decreases. Since the second component of drift vector is the largest, we expect a higher probability for choosing the option 2.

```
>> Cubemodel(500, 3000, 50000, 3, 1, 2, 1, 1, 1, 1)[0]
>> The choice probabilities for option 1, option 2, and option 3
   are 0.242, 0.5333333333333333, and 0.22466666666666665 respectively.
   The mean of the reaction time is 0.35840383474336157.
```

Run another example to see how the results change when $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$:

```
>> Cubemodel(500, 3000, 50000, 3, 0, 0, 0, 3, 1, 1)[0]
>> The choice probabilities for option 1, option 2, and option 3 are
   0.4866666666666667, 0.25233333333333335, and 0.261 respectively.
   The mean of the reaction time is 0.41918838376767537
```

This change impacted the choice probabilities (making option 1 have the highest probability) and decreased the mean of reaction time compared to the first symmetric case. Now consider two cases with a non-zero drift vector:

$$\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{\Sigma}_1 = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \boldsymbol{\Sigma}_2 = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

```
>> Cubemodel(500, 3000, 50000, 3, 1, 1, 1, 3, 1, 1)[0]
>> The choice probabilities for option 1, option 2, and option 3 are
   0.5246666666666666, 0.24766666666666667, and 0.22766666666666666
   respectively. The mean of the reaction time is 0.3073761475229505
```

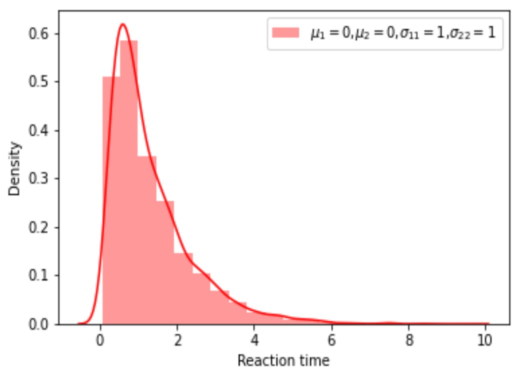
```
>> Cubemodel(500, 3000, 50000, 3, 1, 1, 1, 5, 1, 1)[0]
>> The choice probabilities for option 1, option 2, and option 3 are
   0.526, 0.23866666666666667, and 0.23533333333333334 respectively.
   The mean of the reaction time is 0.27543884211017555
```

We see that even when $\boldsymbol{\mu} \neq \mathbf{0}$ (however has equal components), increasing one of the components of $\boldsymbol{\Sigma}$ decreases the mean of reaction time.

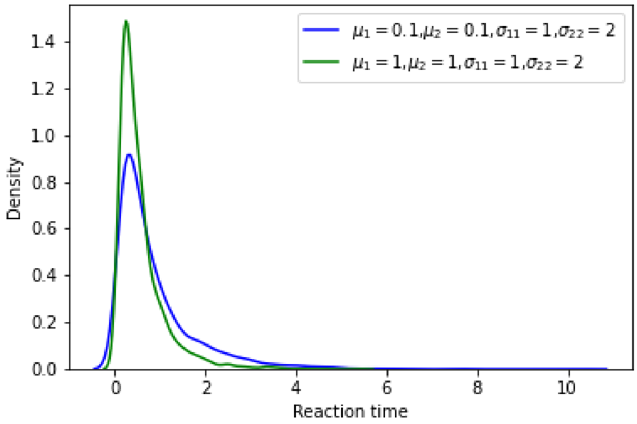
3.4 Reaction time distribution: Cube model with two alternatives and one fall-back option

This part deals with the comparison of the distribution of the reaction time random variable of this model for different $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. The most important implementation tool `Python` offers for this step is `seaborn.distplot` explained in [12]. Intuitively, one should expect that the majority of values is contained in an interval which has the peak in the density curve. To illustrate what we want to show, consider the example whose output is Figure 1:

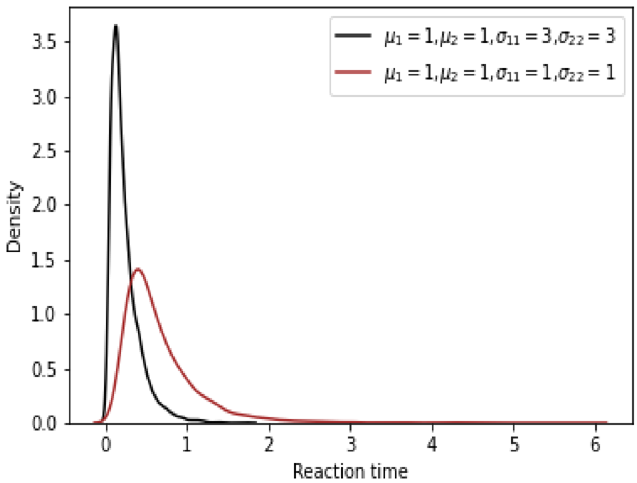
```
>> import seaborn as sns
>> sns.distplot(Cubemodel(100, 3000, 20000, 2, 0, 0, 1, 1), color="r",
   hist=True, norm_hist=True, axlabel="Reaction time",
   label=r"$\mu_{1}=0$,$\mu_{2}=0$,$\sigma_{11}=1$,$\sigma_{22}=1$")
```



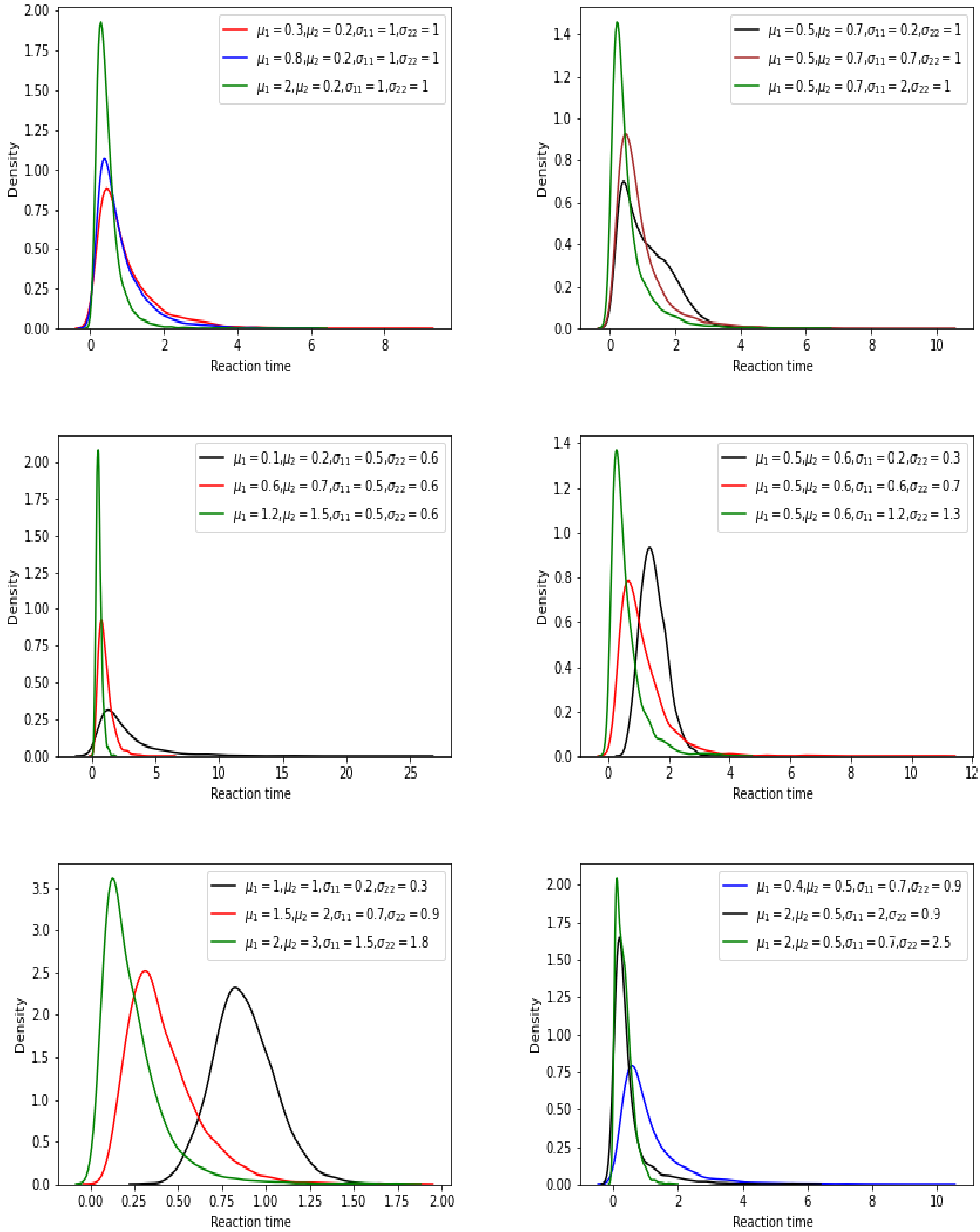
When we increase the magnitude of the drift vector, the density curve gets vertically displaced as below:



The above plot supports the claim that the mean of the reaction time is smaller the larger the magnitude of the drift vector is. For the same number of process simulation, the reaction time values are concentrated in a small interval (green; larger drift), whereas for a smaller drift, they spread out to the right, making the general mean larger (blue). One may wish to observe the effect on the distribution when changing the diagonal entries of the $\mathbf{\Sigma}$:

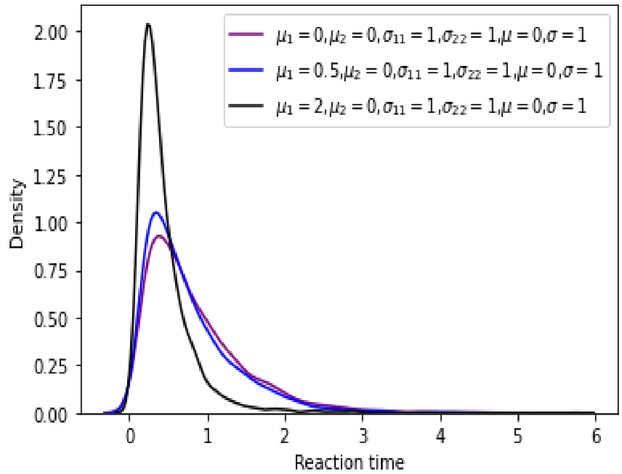


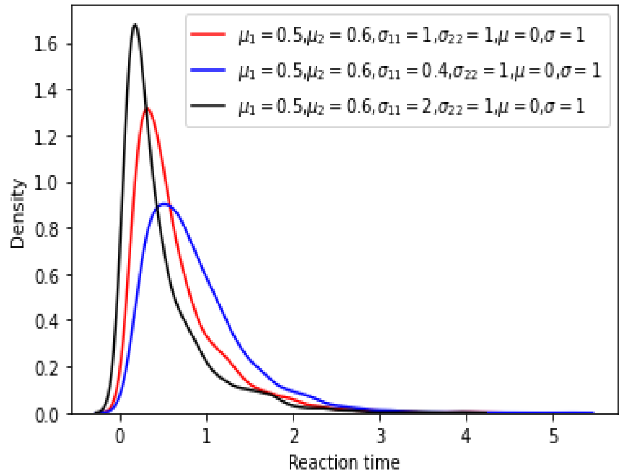
From this figure, we could infer that there is a natural tendency of the reaction time values to decrease when increasing the components of $\mathbf{\Sigma}$. The following pictures display the distributions of reaction times concerning different inputs with 3000 simulations of the process. Their purpose is to convey the change in the distribution when controlling each of the inputs (separately and jointly) of this model.

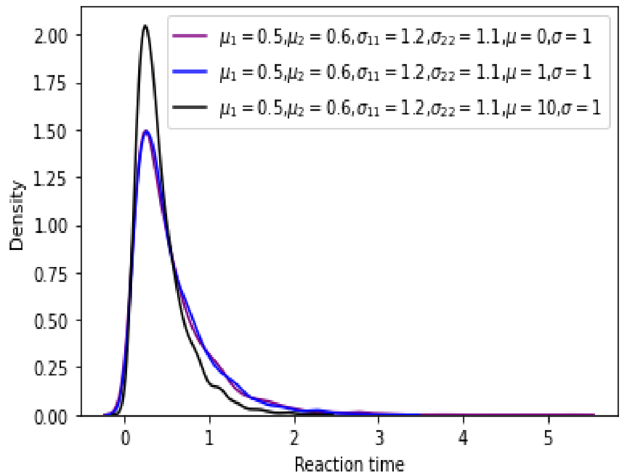

$\mu_1 = 0.3, \mu_2 = 0.2, \sigma_{11} = 1, \sigma_{22} = 1$
$\mu_1 = 0.8, \mu_2 = 0.2, \sigma_{11} = 1, \sigma_{22} = 1$
$\mu_1 = 2, \mu_2 = 0.2, \sigma_{11} = 1, \sigma_{22} = 1$
$\mu_1 = 0.5, \mu_2 = 0.7, \sigma_{11} = 0.2, \sigma_{22} = 1$
$\mu_1 = 0.5, \mu_2 = 0.7, \sigma_{11} = 0.7, \sigma_{22} = 1$
$\mu_1 = 0.5, \mu_2 = 0.7, \sigma_{11} = 2, \sigma_{22} = 1$
$\mu_1 = 0.1, \mu_2 = 0.2, \sigma_{11} = 0.5, \sigma_{22} = 0.6$
$\mu_1 = 0.6, \mu_2 = 0.7, \sigma_{11} = 0.5, \sigma_{22} = 0.6$
$\mu_1 = 1.2, \mu_2 = 1.5, \sigma_{11} = 0.5, \sigma_{22} = 0.6$
$\mu_1 = 0.5, \mu_2 = 0.6, \sigma_{11} = 0.2, \sigma_{22} = 0.3$
$\mu_1 = 0.5, \mu_2 = 0.6, \sigma_{11} = 0.6, \sigma_{22} = 0.7$
$\mu_1 = 0.5, \mu_2 = 0.6, \sigma_{11} = 1.2, \sigma_{22} = 1.3$
$\mu_1 = 1, \mu_2 = 1, \sigma_{11} = 0.2, \sigma_{22} = 0.3$
$\mu_1 = 1.5, \mu_2 = 2, \sigma_{11} = 0.7, \sigma_{22} = 0.9$
$\mu_1 = 2, \mu_2 = 3, \sigma_{11} = 1.5, \sigma_{22} = 1.8$
$\mu_1 = 0.4, \mu_2 = 0.5, \sigma_{11} = 0.7, \sigma_{22} = 0.9$
$\mu_1 = 2, \mu_2 = 0.5, \sigma_{11} = 2, \sigma_{22} = 0.9$
$\mu_1 = 2, \mu_2 = 0.5, \sigma_{11} = 0.7, \sigma_{22} = 2.5$
Density
Reaction time

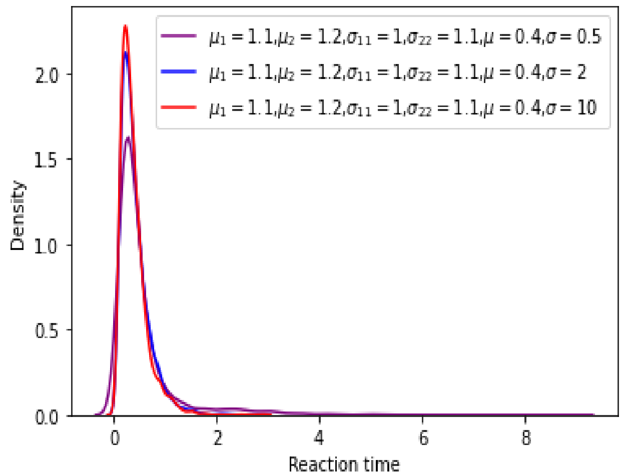
3.5 Reaction time distribution: Disk model

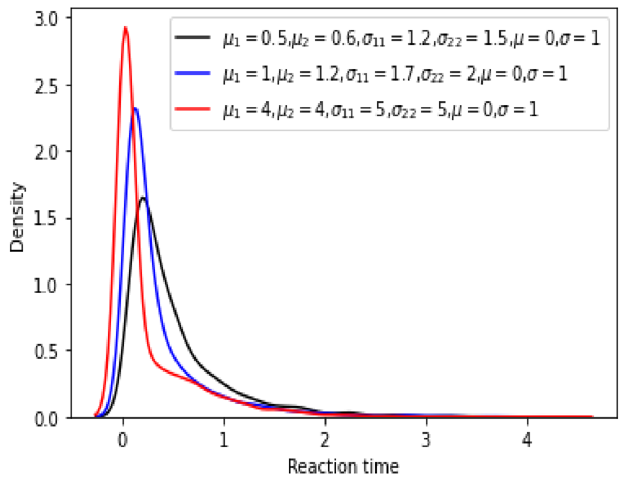
The following plots illustrate the probability distribution of the reaction time when simulating this model 3000 times. In the first four plots, I change only one component of the input variables to observe the resulting change in distribution. In the remaining two plots, I jointly change input variables and obtain the corresponding distributions.

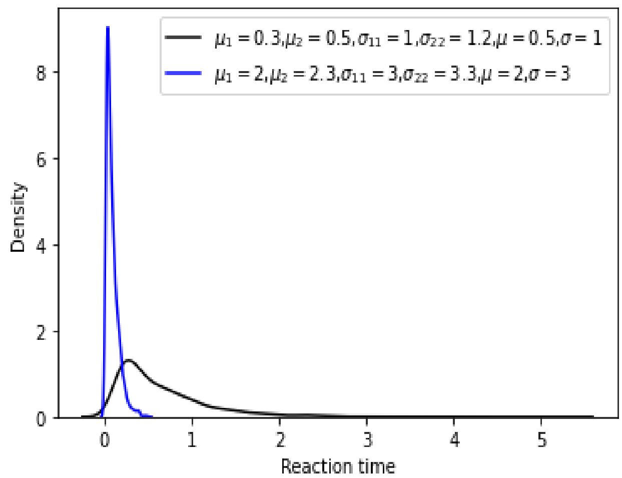






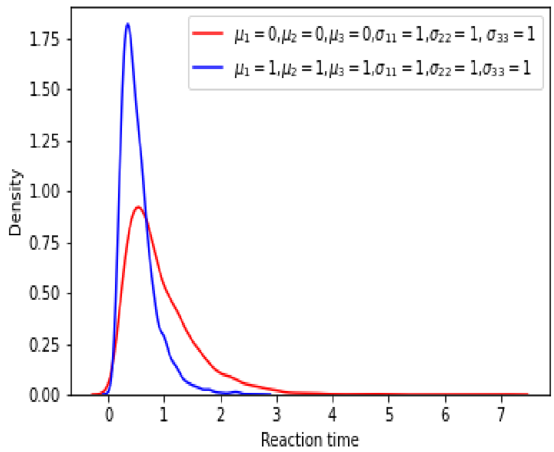


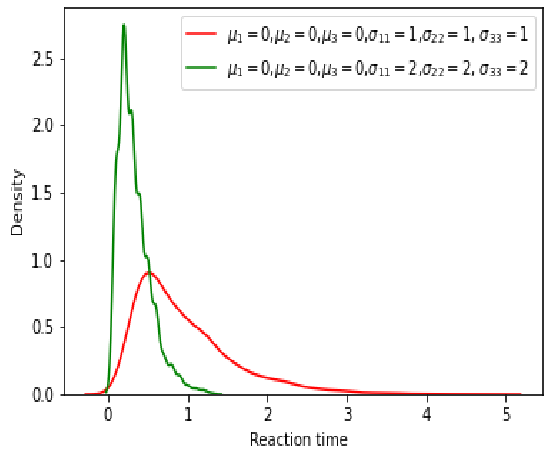


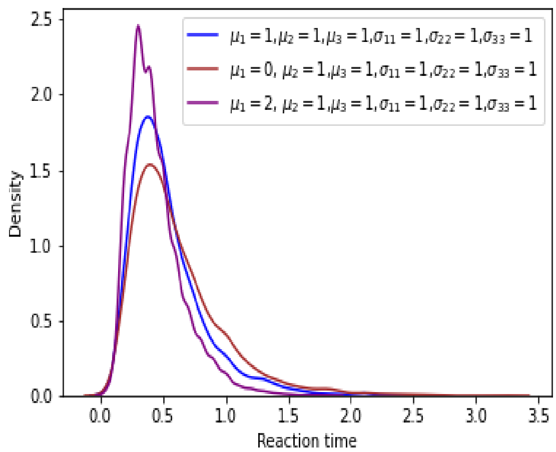


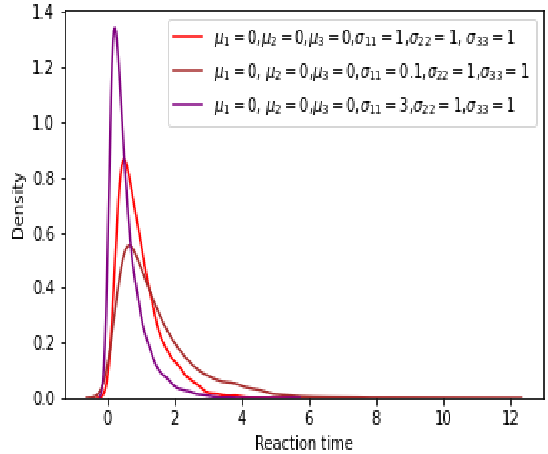
3.6 Reaction time distribution: Cube model with three alternatives

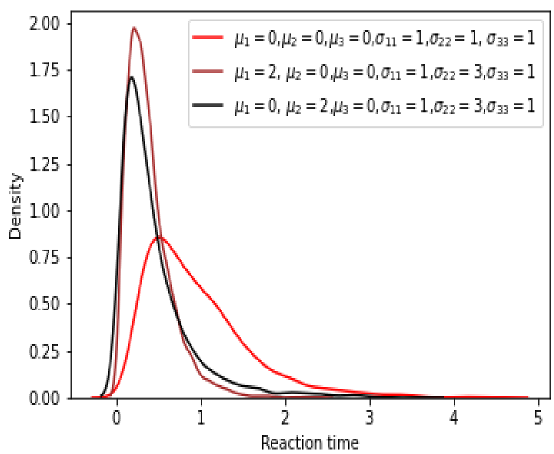
The following pictures demonstrate the change in the probability distribution of the reaction time in this model (3000 simulations), when varying $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$

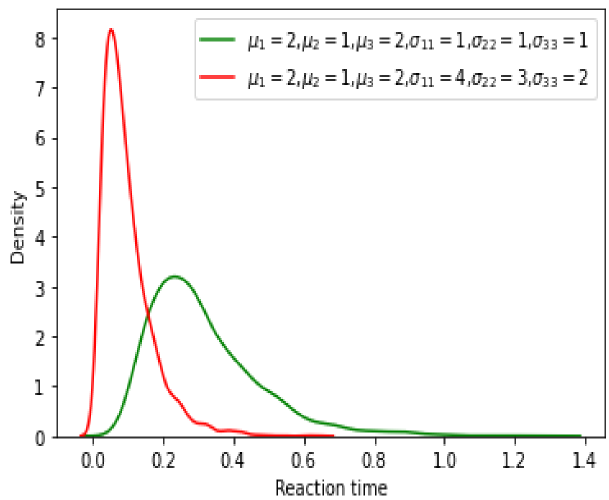








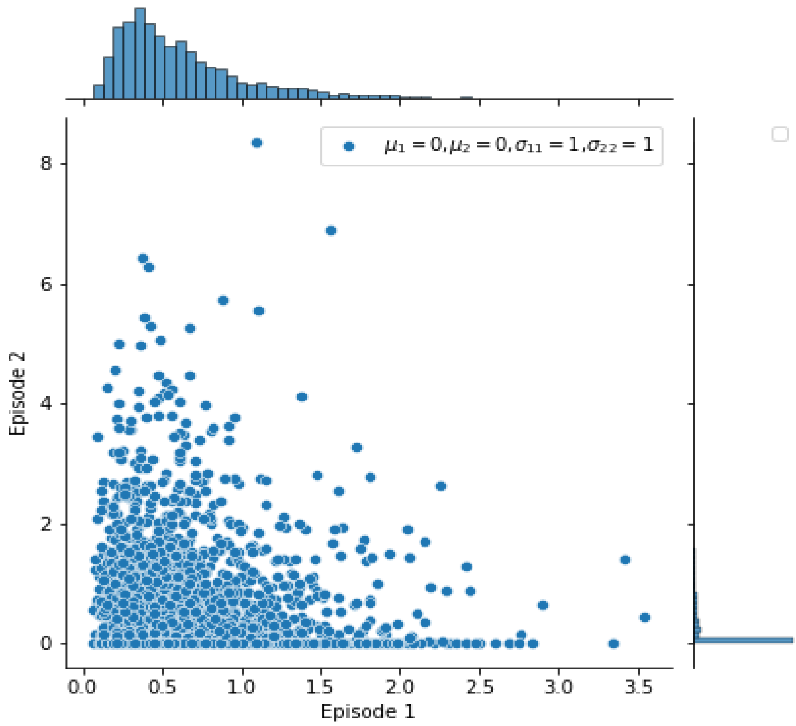




3.7 Episodes: Cube model with two alternatives and one fall-back option

In a setting with two alternatives and one-fall back option, the decision making process of this model terminates in one or two episodes. If the diffusion process initially hits the accept boundary of one of the two available alternatives, the decision making ends in one episode. On the other hand given that the diffusion process initially hits the reject boundary of one of the two options (implying the end of the first episode), it continues until it hits the reject or accept boundary of the remaining option (indicating the end of the second episode and the end of the decision making process). This section illustrates the distribution of the duration of each episode in this model. Let us first see an example of the distributions of these two random variables, implemented in `Python` through `seaborn.jointplot` explained in [13] (3000 simulations, $\boldsymbol{\mu} = \mathbf{0}$, and $\boldsymbol{\Sigma} = \mathbf{I}$):

```
>> sns.jointplot(x="Episode 1",y="Episode 2", data=episodes_new,
   label=r"$\mu_{1}=0$,$\mu_{2}=0$,$\sigma_{11}=1$,$\sigma_{22}=1$")

>>
```



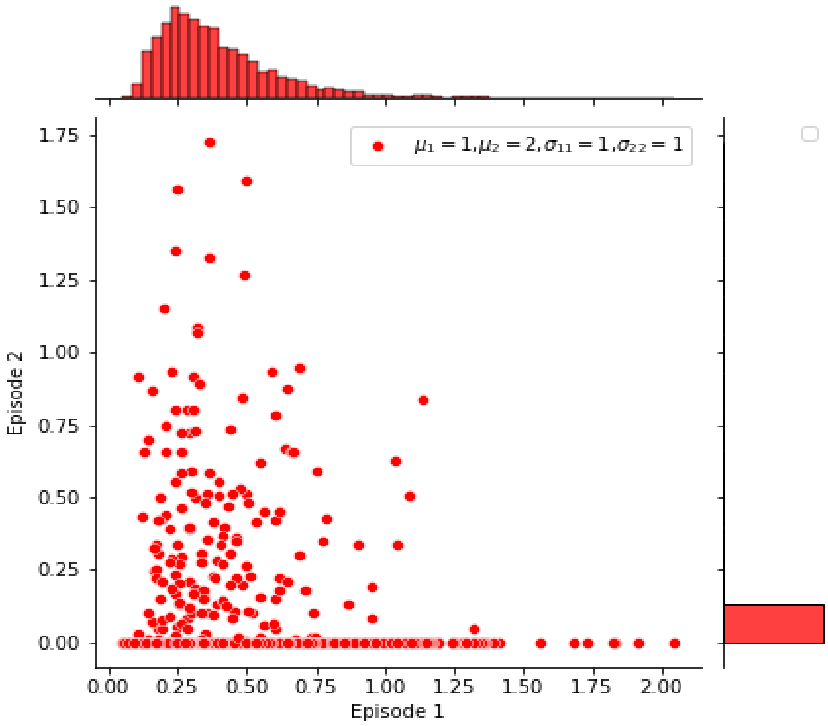
If the process ends in one episode, the duration for the second episode is set to be 0. This plot shows not only the individual distributions through the histograms, but also the joint occurrence of the duration for each episode. Intuitively we should expect bell curves for each of the durations, similarly to the reaction time distribution. To obtain quantitative insights about each of the episodes, we run the code that finds the mean of the time corresponding to the duration of each episode:

```
>> Cubemodel(100, 3000, 20000, 2, 0, 0, 1, 1)[4]
>> The average durations of the first and second episode of the
   process are 0.6282464123206161 and 0.4493358001233395 respectively
```

Let us analyze how the increase of the magnitude of the drift vector influences the duration of the first and second episode. Choose $\boldsymbol{\mu} = (1, 2)^\top$ and $\boldsymbol{\Sigma} = \mathbf{I}$:

```
>> Cubemodel(100, 3000, 20000, 2, 1, 2, 1, 1)[4]
>> The average durations of the first and second episode of the
   process are 0.39619647649049117 and 0.024774572061936426 respectively
```

We observe that the average durations of each episode decrease, which is not a surprise, since the average stopping time of the whole process decreases when the magnitude of $\boldsymbol{\mu}$ increases. However, it is interesting to note that the average duration of the second episode is much smaller. This suggests that the process with these parameters, ends mostly in one episode. The following plot supports this claim; mainly, most of the 3000 data points (simulations) do not have a second episode (Episode 2=0):



From our previous simulations, we know that using this drift vector implies the highest probability for choosing option 2. This can indicate that the process hits mostly the accept boundary of option 2, in the very first episode.

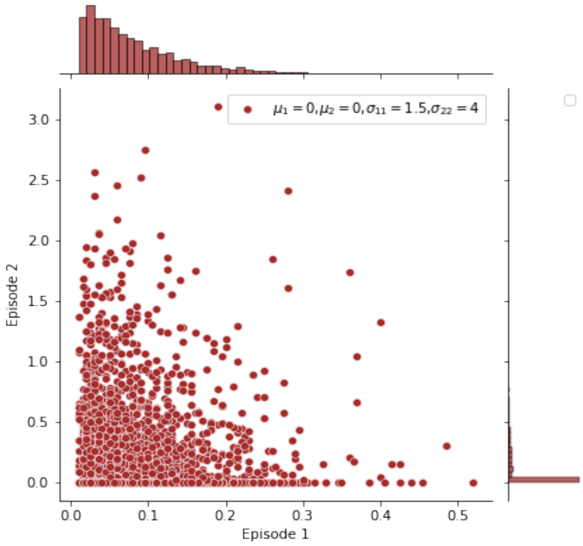
Let us now see how the increase of $\boldsymbol{\Sigma}$ components impacts the results when $\boldsymbol{\mu} = \mathbf{0}$. Choose the components of it to be $\sigma_{11} = 1.5$ and $\sigma_{22} = 2.5$:

```
>> Cubemodel(100, 3000, 20000, 2, 0, 0, 1.5, 2.5)[4]
>> The average durations of the first and second episode of the
   process are 0.1675400436688501 and 0.188104405220261 respectively.
```

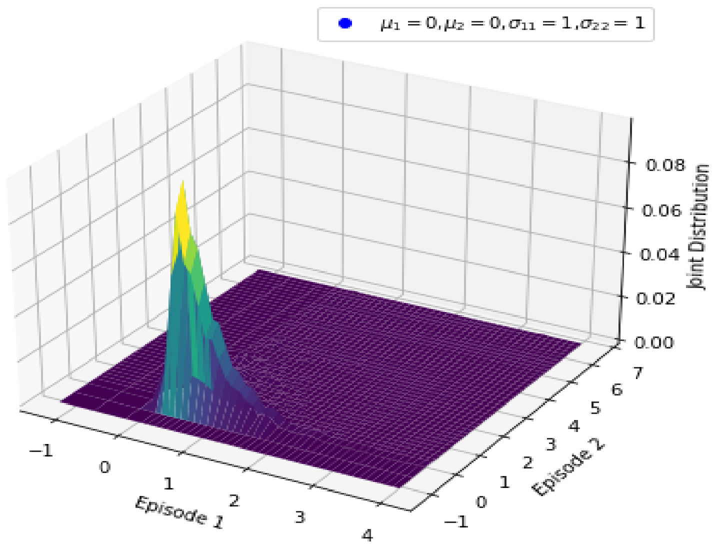
Take another example where the components of $\mathbf{\Sigma}$ vary more i.e., consider $\sigma_{11} = 1.5$ and $\sigma_{22} = 4$.

```
>> Cubemodel(100, 3000, 20000, 2, 0, 0, 1.5, 4)[4]
>> The average durations of the first and second episode of the
   process are 0.08328083070820208 and 0.21289564478223913 respectively.
```

We observe that the durations of each episode is much smaller compared to the results in the very first case. Let us plot the results:

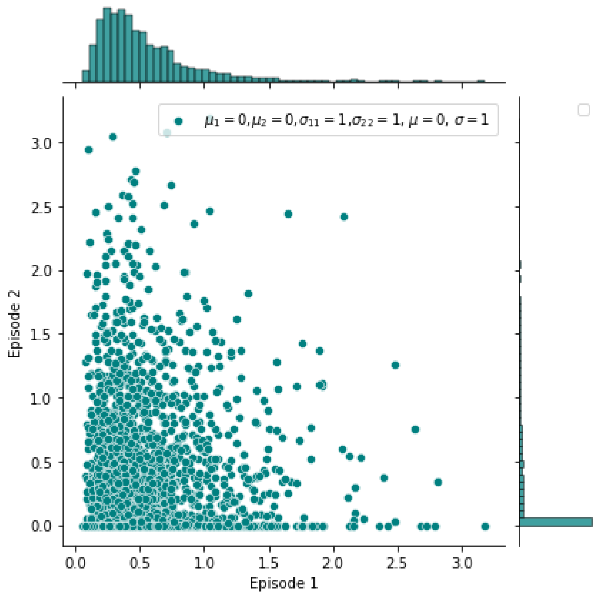


The joint distribution of the durations of each episode is not implemented by a built-in `Python` function, but through algorithmic code lines implemented by myself that can be found in the **Appendix**. The displaying function in `Python` used to obtain the following plot is `plot_surface` explained in [14]. The following plot is an example of a joint distribution for $\boldsymbol{\mu} = 0$ and $\mathbf{\Sigma} = \mathbf{I}$.



3.8 Episodes: Disk model

Disk model describes the decision making setting with two alternatives and one fall-back option. Similarly to the previous Cube model, it implies that the decision making process terminates in one or two episodes (the four boundaries of this model belong to the unit disk), I consider the time step to be 0.005 and simulated the process 3000 times to obtain each of the following results. Let us obtain the distributions of the durations of each episode as well as their mean value, for the typical example where $\boldsymbol{\mu} = \mathbf{0}$, $\boldsymbol{\Sigma} = \mathbf{I}$, μ=0, and $\sigma = 1$:

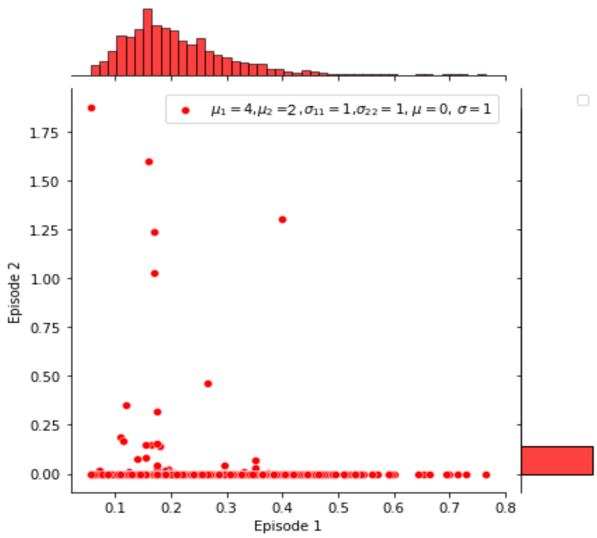



```
>> Diskmodel(100, 3000, 20000, 0, 0, 1, 1, 0, 1)[4]
>> The average durations of the first and second episode of
   the process are 0.5339816990849543 and 0.28009567145023917
   respectively.
```

Proposition 2 in the first chapter, implies that mean of the duration for the first episode in this setting is 0.5. In the above output, we can see that the simulation result is quite near the true value. However, if we try smaller time steps and more simulations of the process, we can increase the accuracy of that implemented value of Episode 1. Let us choose $\boldsymbol{\mu} = (4, 2)^\top$ and $\boldsymbol{\Sigma} = \mathbf{I}$, $\mu = 0$, and $\sigma = 1$:

```
>> Diskmodel(100, 3000, 20000, 4, 2, 1, 1, 0, 1)[4]
>> The average durations of the first and second episode of
   the process are 0.20807873727019682 and 0.003188492757971232
   respectively.
```

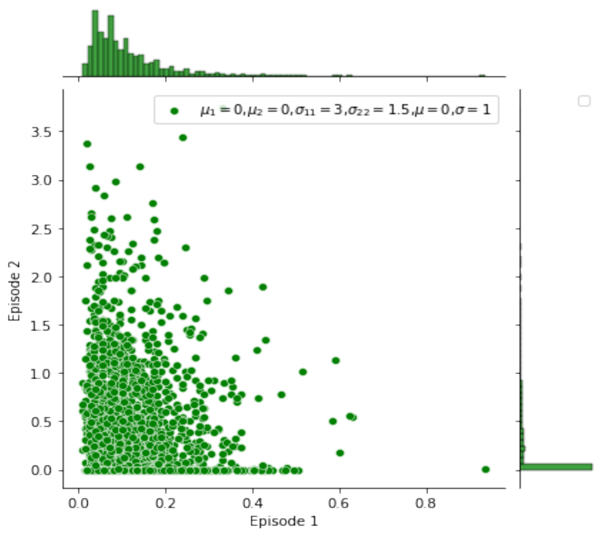
These results imply that most simulations of the process will end in one episode (since the mean of duration of the second episode is negligible, suggesting that most values equal 0). Similarly to the previous model, we could infer that this drift vector indicates that option 1 has the highest probability of being chosen. Combining these two observations, we conclude that the majority of simulations will hit the accept boundary of option 1 for the very first time indicating the end of decision making in one episode. The following plot supports our claim concerning the duration of the second episode.



Let us increase the components of $\mathbf{\Sigma}$ in the absence of a drift vector and observe the results. Take $\sigma_{11} = 3$, and $\sigma_{22} = 1.5$:

```
>> Diskmodel(100, 3000, 20000, 0, 0, 3, 1.5, 0, 1)
>> The average durations of the first and second episode of
   the process are 0.10927879727319699 and 0.3025451272563628
   respectively.
```

Proposition 2 suggests that for this drift vector, μ, and σ, increasing the norm of $\mathbf{\Sigma}$ implies a smaller mean of duration of the first episode. When compared to the example where $\mathbf{\Sigma} = \mathbf{I}$, the above simulation result supports this claim by showing that the mean duration of the first episode decreases. One may wish to see their joint occurrence through the following plot:



We can increase μ and observe how the duration of each episode changes. Let us run this model twice with the following input parameters (changing only μ):

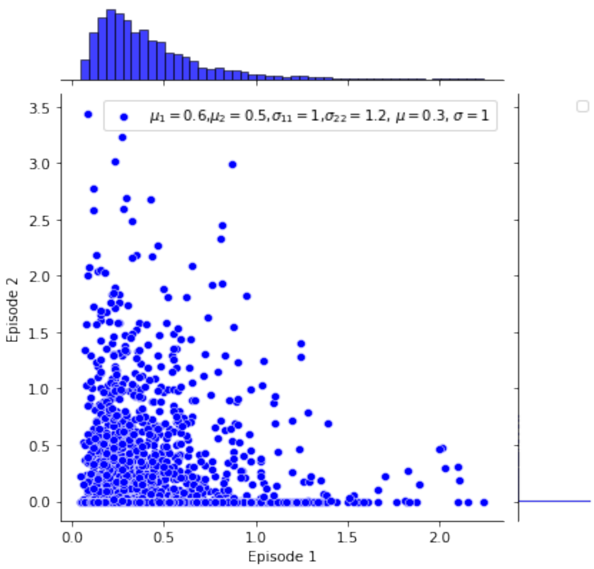
$$\boldsymbol{\mu} = \begin{pmatrix} 0.6 \\ 0.5 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 1.2 \end{pmatrix}, \quad \mu_1 = 0.3, \quad \mu_2 = 2.5, \quad \sigma = 1.$$

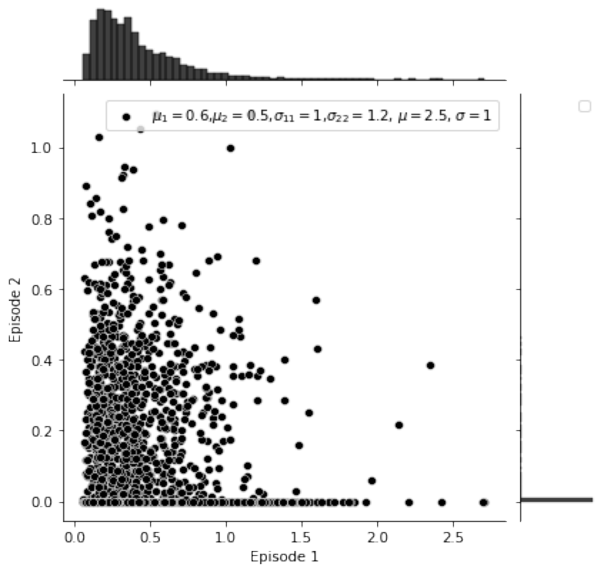
Given that only the second episode of the process is governed by the circular Wiener process as in (1), we should expect that the increase of μ implies the decrease of the average duration of the second episode. Run the code for each of these two cases and obtain the results:

```
>> Diskmodel(100, 3000, 20000, 0.6, 0.5, 1, 1.2, 0.3, 1)[4]
>> The average durations the first and second episode of
   the process are 0.4219094288047736 and 0.15724786239311966
   respectively
```

```
>> Diskmodel(100, 3000, 20000, 0.6, 0.5, 1, 1.2, 2.5, 1)[4]
>> The average durations of the first and second episode of
   the process are 0.4194726402986816 and 0.08038401920096004
   respectively.
```

The duration of the first episode seem to not change in both cases (which should be true since the diffusion process in the first episode is independent of μ). However, the mean decreases in the second episode (for a larger μ), which supports our claim. The plots demonstrate the individual distributions of the durations relevant to each episode:



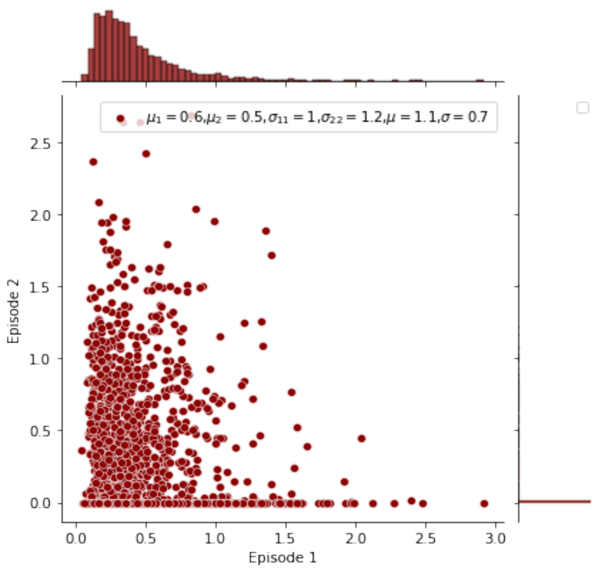


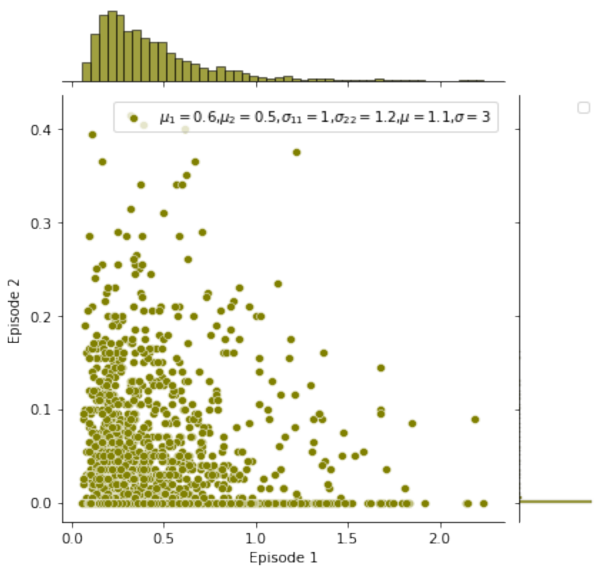
Let us observe the effect of increasing σ in the duration of each episode. Similarly, we expect that only the duration of second episode will be influence by this change. Run the code with the following inputs:

$$\boldsymbol{\mu} = \begin{pmatrix} 0.6 \\ 0.5 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 1.2 \end{pmatrix}, \quad \mu = 1.1, \quad \sigma_1 = 0.7, \quad \sigma_2 = 3.$$

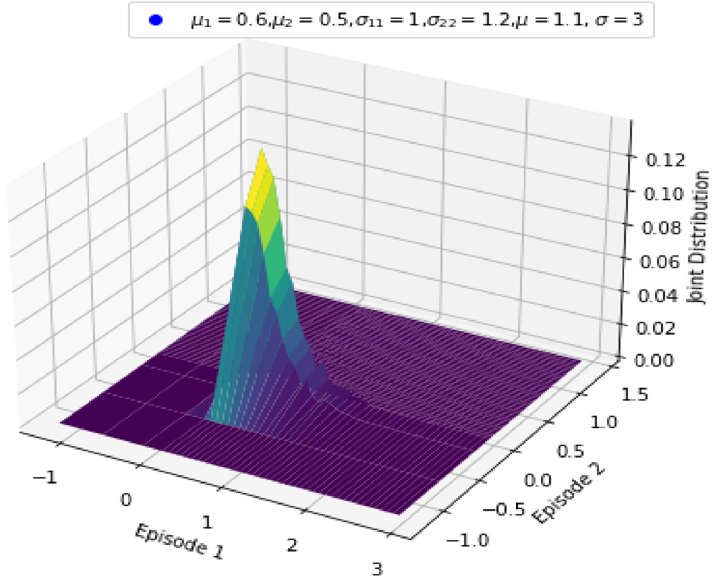
```
>> Diskmodel(100, 3000, 20000, 0.6, 0.5, 1, 1.2, 1.1, 0.7)[4]
>> The average durations of the first and second episode
   of the process are 0.4239795323099488 and 0.17030851542577127
   respectively

>> Diskmodel(100, 3000, 20000, 0.6, 0.5, 1, 1.2, 1.1, 3)[4]
>> The average durations of the first and second episode
   of the process are 0.4272296948180742 and 0.02281140557027785
   respectively
```

This result should be a surprise, because in the distribution of reaction time of this model, we observed that when increasing only the value of σ, the plots moved up vertically, indicating a lower mean of the reaction time. Since the increase of this input influences only the duration of the second episode, we should expect a lower mean of this duration. The plots below show the distribution of the duration of each episode, for these two cases.





Finally, let us close the analysis of the duration of episodes for this model, by showing an example of the joint distribution of these two random variables. The input values are indicated in the legend.



3.9 Episodes: Cube model with three alternatives

The Cube model with three alternatives implies that the decision making process ends in one or two episodes. If the 3-dimensional diffusion process hits first the accept boundary of one of the options, we say that the process ends in one episode. However, if it initially hits one reject boundary, the process continues as a 2-dimensional diffusion process. The second episode ends as soon as a reject or accept boundary of the remaining options is hit. If, in this case, a reject boundary is hit first, the process ends with the choice of the last remaining option. Let us analyze how the duration of each episode changes, when increasing one of the components of the drift $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Consider the first example to have the inputs:

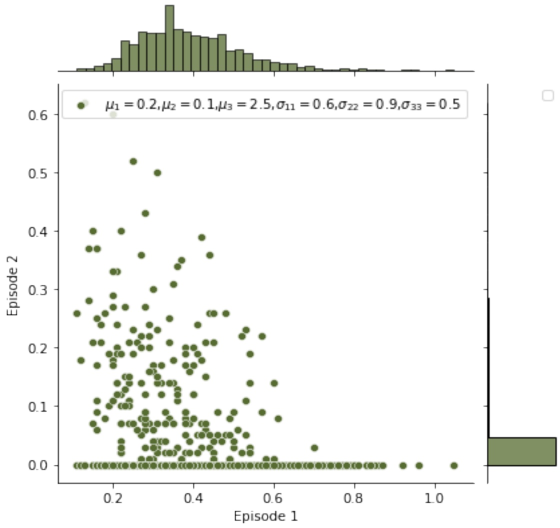
$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{pmatrix} = \begin{pmatrix} 0.2 \\ 0.1 \\ 0.3 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} = \begin{pmatrix} 0.6 & 0 & 0 \\ 0 & 0.9 & 0 \\ 0 & 0 & 0.5 \end{pmatrix},$$

Simulating the process 3000 times and using a time step 0.01, we obtain the following results:

```
>> Cubemodel(500, 3000, 50000, 3, 0.2, 0.1, 0.3, 0.6, 0.9, 0.5)[4]
>> The average durations of the first and second episode of the process
   are 0.9552991059821198 and 0.3647639619459056 respectively
```

Choose $\mu_3 = 2.5$ and keep all the other parameters unchanged. Obtain the means of durations and the plot as below :

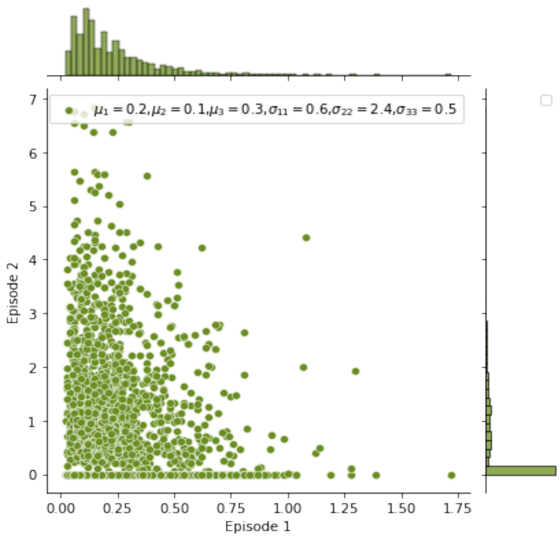
```
>> Cubemodel(500, 3000, 50000, 3, 0.2, 0.1, 2.5, 0.6, 0.9, 0.5)[4]
>> The average durations the first and second episode of the process
   are 0.3897011273558805 and 0.009366854003746743 respectively.
```



The choice of the new drift vector implies the higher probability of choosing option 3. The plot indicates that the majority of simulations end in one episode. We can conclude that the majority of these simulations hit the accept boundary of option 3 for the first time.

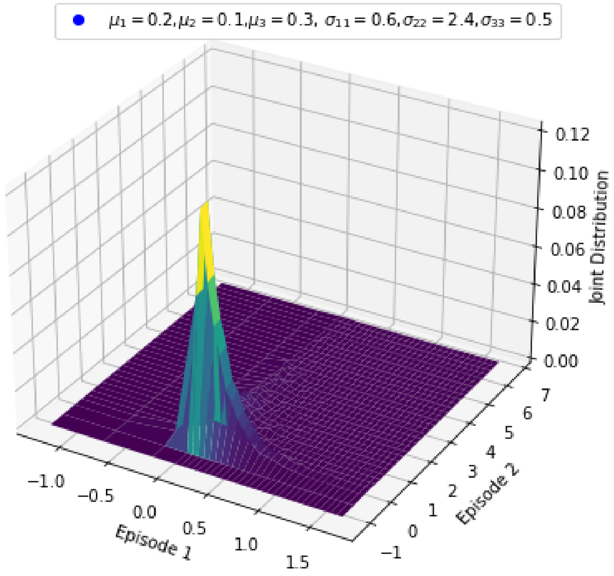
Secondly, take $\sigma_{22} = 2.4$ and let the other inputs be as in the first example. We get:

```
>> Cubemodel(500, 3000, 50000, 3, 0.2, 0.1, 0.3, 0.6, 2.4, 0.5)[4]
>> The average durations of the first and second episode of the process
   are 0.2246911604898765 and 0.7192543850877019 respectively.
```



We can infer from the picture that the process does not end in one single episode in most simulations. This suggests that the reject boundaries have the tendency to be hit first.

We close this section by showing the joint distribution of the time random variables determining the end of the first and second episode of this model with the above indicated parameters:



3.10 Worst-best option ranking: Cube model with two alternatives as one fall-back option

Apart from obtaining fundamental results from this model, one could aim to obtain new results which motivate further research on this model. This section finds the probabilities for settings describing the degree of preference towards each of three alternatives in this model, These settings imply the formation of new events which express each of the alternatives as being the most/least favorite one. The results of this section include the probabilities for each of these events and time needed to complete them. Let the set of alternatives one chooses from be $\{A_1, A_2, F\}$ where F is the fall-back option. We obtain the following settings (they can be obtained similarly in the Disk model):

A_1A_2F : The two dimensional diffusion process hits the accept boundary of A_1 first. The process continues as a one-dimensional diffusion process, hitting the accept boundary of A_2 initially. This suggests that A_1 is more desirable than A_2 and A_2 more desirable that F.

A_1FA_2 : The accept boundary of A_1 is hit initially. This makes option A_1 the most desirable option. The process continues and hits for the first time the reject boundary of A_2. This makes option A_2 the least preferred option. OR. The reject boundary of A_2 is hit for the first time. The process continues a second episode in which it hits the accept boundary of A_1 first.

A_2A_1F : The accept boundary of A_2 is hit for the first time. The process continues and hits the accept boundary of A_1 first.

A_2FA_1 : The accept boundary of A_2 is hit firstly. The process in the second episode hits the reject boundary of A_1 initially. OR. The reject boundary of A_1 is hit first. The process continues a second episode and hits for the first time the accept boundary of A_2.

FA_1A_2 : The reject boundary of A_2 is hit first. This makes option A_2 the least preferred one. The process continues and hits the reject boundary of option A_1. Excluding these two options, makes option F the most desirable one.

FA_2A_1 : The reject boundary of A_1 is hit for the first time. The one dimensional diffusion process of the second episode hits the reject boundary of option A_2 initially.

The following table provides the combination of these scenarios that are associated to events of interest for this model:

Table 1: The most and least preferred option

Scenarios	Events
$A_1A_2F \cup A_1FA_2$	A_1 is the most preferred option = A_1^+
$A_2FA_1 \cup FA_2A_1$	A_1 is the least preferred option = A_1^-
$A_2A_1F \cup A_2FA_1$	A_2 is the most preferred option = A_2^+
$A_1FA_2 \cup FA_1A_2$	A_2 is the least preferred option = A_2^-
$FA_1A_2 \cup FA_2A_1$	The fall-back option (F) is the most preferred = F^+
$A_1A_2F \cup A_2A_1F$	F is the least preferred option = F^-

To illustrate the relevant results of this model, consider the example with $\boldsymbol{\mu} = \mathbf{0}$, $\Sigma = \mathbf{I}$, 3000 simulations, and a time step of 0.005:

Table 2: The probabilities for each of the settings

$P[A_1A_2F]$	$P[A_1FA_2]$	$P[A_2A_1F]$	$P[A_2FA_1]$	$P[FA_1A_2]$	$P[FA_2A_1]$
0.116667	0.258333	0.124	0.259333	0.123667	0.118

After running this code multiple times, it is worth noting that values of $P[A_2FA_2]$ and $P[A_2FA_1]$ do not change too much, but stay near 0.25. This fact may motivate a conjecture. The probabilities for each of the events in Table 1 are displayed in Table 3.

Table 3: The probabilities for each of the events

$P[A_1^+] = P[A_1A_2F] + P[A_1FA_2]$	$P[A_2^+] = P[A_2A_1F] + P[A_2FA_1]$	$P[F^+] = P[FA_1A_2] + P[FA_2A_1]$
0.375	0.383333	0.241667
$P[A_1^-] = P[A_2FA_1] + P[FA_2A_1]$	$P[A_2^-] = P[A_1FA_2] + P[FA_1A_2]$	$P[F^-] = P[A_1A_2F] + P[A_2A_1F]$
0.377333	0.382	0.240667

The means of the time needed to complete these events are shown in Table 4:

Table 4: Mean time for the events

A_1^+	A_1^-	A_2^+	A_2^-	F^+	F^-
0.937789	0.945091	0.937789	0.926735	1.588438	1.57316

The results obtained are very symmetric with respect to every option (e.g., the probabilities of A_1^+ and A_1^- nearly coincide). This follows from the symmetric input values we considered for $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. The following tables shows the results for different inputs.

Table 5: $\mu_1 = 1.5, \mu_2 = 1.1, \sigma_{11} = 1, \sigma_{22} = 1.2$

$P[A_1A_2F]$	$P[A_1FA_2]$	$P[A_2A_1F]$	$P[A_2FA_1]$	$P[FA_1A_2]$	$P[FA_2A_1]$
0.372333	0.178333	0.401667	0.039	0.003667	0.005
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[F^+]$	$P[F^-]$
0.550667	0.044	0.440667	0.182	0.008667	0.774
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	F^+/mean time	F^-/mean time
0.493372	0.683936	0.39359	0.610543	0.842542	0.868639

Table 6: $\mu_1 = 3, \mu_2 = 1.1, \sigma_{11} = 1, \sigma_{22} = 1.2$

$P[A_1A_2F]$	$P[A_1FA_2]$	$P[A_2A_1F]$	$P[A_2FA_1]$	$P[FA_1A_2]$	$P[FA_2A_1]$
0.570333	0.165	0.263333	0.001333	0.0	0.0
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[F^+]$	$P[F^-]$
0.735333	0.001333	0.264667	0.165	0.0	0.833667
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	F^+/mean time	F^-/mean time
0.301046	0.535027	0.275448	0.638739	0.255013	0.67291

Table 7: $\mu_1 = 1.5, \mu_2 = 3, \sigma_{11} = 1, \sigma_{22} = 1.2$

$P[A_1A_2F]$	$P[A_1FA_2]$	$P[A_2A_1F]$	$P[A_2FA_1]$	$P[FA_1A_2]$	$P[FA_2A_1]$
0.255333	0.011667	0.691667	0.041	0.0	0.000333
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[F^+]$	$P[F^-]$
0.267	0.041333	0.732667	0.011667	0.000333	0.947
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	F^+/mean time	F^-/mean time
0.329267	0.666525	0.27935	0.302015	0.515026	0.708916

Table 8: $\mu_1 = 1.5, \mu_2 = 1.1, \sigma_{11} = 3, \sigma_{22} = 1.2$

$P[A_1A_2F]$	$P[A_1FA_2]$	$P[A_2A_1F]$	$P[A_2FA_1]$	$P[FA_1A_2]$	$P[FA_2A_1]$
0.469333	0.084333	0.038667	0.332	0.007667	0.068
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[F^+]$	$P[F^-]$
0.553667	0.4	0.370667	0.092	0.075667	0.508
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	F^+/mean time	F^-/mean time
0.128742	0.129661	0.60311	0.595809	0.648975	0.618984

Table 9: $\mu_1 = 1.5, \mu_2 = 1.1, \sigma_{11} = 1, \sigma_{22} = 3$

$P[A_1A_2F]$	$P[A_1FA_2]$	$P[A_2A_1F]$	$P[A_2FA_1]$	$P[FA_1A_2]$	$P[FA_2A_1]$
0.035333	0.407667	0.512667	0.026333	0.017	0.001
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[F^+]$	$P[F^-]$
0.443	0.027333	0.539	0.424667	0.018	0.548
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	F^+/mean time	F^-/mean time
0.596396	0.678205	0.125111	0.134316	0.665311	0.626537

3.11 Worst-best option ranking: Disk model

The same scenarios and events apply for this model (in this case we have the disk boundary and the accumulation process of the second episode is the circular Wiener process). We simulate the process 3000 times with a time step 0.005 and obtain the following results for different input values (μ_1, μ_2, σ_{11}, σ_{22}, μ, and σ):

Table 10: $\mu_1 = 0, \mu_2 = 0, \sigma_{11} = 1, \sigma_{22} = 1, \mu = 0, \sigma = 1$

$P[A_1A_2F]$	$P[A_1FA_2]$	$P[A_2A_1F]$	$P[A_2FA_1]$	$P[FA_1A_2]$	$P[FA_2A_1]$
0.128	0.259	0.115	0.263667	0.125	0.109333
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[F^+]$	$P[F^-]$
0.387	0.373	0.378667	0.384	0.234333	0.243
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	F^+/mean time	F^-/mean time
0.717848	0.716814	0.735032	0.743349	1.124068	1.064285

Table 11: $\mu_1 = 1.7, \mu_2 = 1, \sigma_{11} = 1.1, \sigma_{22} = 1.2, \mu = 1.3, \sigma = 1$

$P[A_1A_2F]$	$P[A_1FA_2]$	$P[A_2A_1F]$	$P[A_2FA_1]$	$P[FA_1A_2]$	$P[FA_2A_1]$
0.437667	0.169667	0.102	0.252	0.010333	0.028333
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[F^+]$	$P[F^-]$
0.607333	0.280333	0.354	0.18	0.038667	0.539667
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	F^+/mean time	F^-/mean time
0.379039	0.766311	0.317568	0.483672	0.755081	0.698692

Table 12: $\mu_1 = 3.3, \mu_2 = 1, \sigma_{11} = 1.1, \sigma_{22} = 1.2, \mu = 1.3, \sigma = 1$

$P[A_1A_2F]$	$P[A_1FA_2]$	$P[A_2A_1F]$	$P[A_2FA_1]$	$P[FA_1A_2]$	$P[FA_2A_1]$
0.593333	0.159333	0.094667	0.145333	0.004667	0.002667
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[F^+]$	$P[F^-]$
0.752667	0.148	0.24	0.164	0.007333	0.688
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	F^+/mean time	F^-/mean time
0.253824	0.842621	0.230817	0.452136	0.55571	0.601898

Table 13: $\mu_1 = 1.7, \mu_2 = 1, \sigma_{11} = 3.2, \sigma_{22} = 2, \mu = 1.3, \sigma = 1$

$P[A_1A_2F]$	$P[A_1FA_2]$	$P[A_2A_1F]$	$P[A_2FA_1]$	$P[FA_1A_2]$	$P[FA_2A_1]$
0.362	0.164667	0.029667	0.166333	0.022333	0.255
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[F^+]$	$P[F^-]$
0.526667	0.421333	0.196	0.187	0.277333	0.391667
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	F^+/mean time	F^-/mean time
0.139244	0.194883	0.211159	0.290995	0.537028	0.487484

Table 14: $\mu_1 = 1.7, \mu_2 = 1, \sigma_{11} = 1.1, \sigma_{22} = 1.2, \mu = 2.5, \sigma = 2.2$

$P[A_1A_2F]$	$P[A_1FA_2]$	$P[A_2A_1F]$	$P[A_2FA_1]$	$P[FA_1A_2]$	$P[FA_2A_1]$
0.362333	0.22	0.177333	0.202	0.019667	0.018667
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[F^+]$	$P[F^-]$
0.582333	0.220667	0.379333	0.239667	0.038333	0.539667
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	F^+/mean time	F^-/mean time
0.351866	0.457651	0.322819	0.408233	0.553289	0.444701

3.12 Worst-best option ranking: Cube model with three alternatives

The settings and events for this model will be slightly more different, since we don't have a fall-back option. The set of alternatives one can choose from is $\{A_1, A_2, A_3\}$. The scenarios expressing the degree of preference towards each of these options involve two possibilities for the outcome of the three dimensional diffusion process in the first episode. These scenarios are:

$A_1A_2A_3$: The three dimensional diffusion process hits the accept boundary of A_1 initially. The process continues as a two dimensional diffusion process (with respect to A_2 and A_3) and hits the accept boundary of A_2 or the reject boundary of A_3 first. OR. The three dimensional diffusion process, hits the reject boundary of A_3 for the first time, The process continues as a two dimensional diffusion process hitting either the accept boundary of A_1 or the reject boundary of A_2 initially.

$A_1A_3A_2$: The process hits the accept boundary of A_1 for the first time, The process continues and hits the accept boundary of A_3 or the reject boundary of A_2 initially. OR. The process hits the reject boundary of A_2 for the first time, The process continues and hits the accept boundary of A_1 or the reject boundary of A_3 initially.

$A_2A_1A_3$: The process hits the accept boundary of A_2 initially. The process in the second episode hits the accept boundary of A_1 or the reject boundary of A_3 first. OR. The process hits the reject boundary of A_3 initially. The process in the second episode hits the accept boundary of A_2 or the reject boundary of A_1 first.

$A_2A_3A_1$: The accept boundary of A_2 is hit for the first time. The process continues and hits the accept boundary of A_3 or the reject boundary of A_1 initially. OR. The reject boundary of A_1 is hit for the first time. The process continues and hits the accept boundary of A_2 or the reject boundary of A_3 first.

$A_3A_1A_2$: The accept boundary of A_3 is hit first. The process in the second episode hits the accept boundary of A_1 or the reject boundary of A_2 initially. OR. The reject boundary of A_2 is hit first. The process in the second episode hits the accept boundary of A_3 or the reject boundary of A_1 initially.

$A_3A_2A_1$: The process hits the accept boundary of A_3 first. The process of the second episode hits the accept boundary of A_2 or the reject boundary of A_1 for the first time. OR. The process hits the reject boundary of A_1 first. The process of second episode hits the accept boundary of A_3 or the reject boundary of A_2 for the first time.

Pairwise combinations of some scenarios generate the events of interest for this model:

Table 15: The most and least preferred option

Scenarios	Events
$A_1A_2A_3 \cup A_1A_3A_2$	A_1 is the most preferred option = A_1^+
$A_2A_3A_1 \cup A_3A_2A_1$	A_1 is the least preferred option = A_1^-
$A_2A_1A_3 \cup A_2A_3A_1$	A_2 is the most preferred option = A_2^+
$A_1A_3A_2 \cup A_3A_1A_2$	A_2 is the least preferred option = A_2^-
$A_3A_1A_2 \cup A_3A_2A_1$	A_3 is the most preferred option = A_3^+
$A_1A_2A_3 \cup A_2A_1A_3$	A_3 is the least preferred option = A_3^-

The following tables with the corresponding input parameters (3000 simulations 0.01 time step), display the choice probabilities for the scenarios, events, and mean of the time needed to accomplish each of the events.

Table 16: $\mu_1 = 0, \mu_2 = 0, \mu_3 = 0, \sigma_{11} = 1, \sigma_{22} = 1, \sigma_{33} = 1$

$P[A_1A_2A_3]$	$P[A_1A_3A_2]$	$P[A_2A_1A_3]$	$P[A_2A_3A_1]$	$P[A_3A_1A_2]$	$P[A_3A_2A_1]$
0.166333	0.181667	0.161667	0.169333	0.164667	0.156333
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[A_3^+]$	$P[A_3^-]$
0.348	0.325667	0.331	0.346333	0.321	0.328
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	A_3^+/mean time	A_3^-/mean time
0.74944	0.348	0.73079	0.331	0.747938	0.321

Table 17: $\mu_1 = 1.2, \mu_2 = 1, \mu_3 = 1.1, \sigma_{11} = 1.1, \sigma_{22} = 0.9, \sigma_{33} = 1.2$

$P[A_1A_2A_3]$	$P[A_1A_3A_2]$	$P[A_2A_1A_3]$	$P[A_2A_3A_1]$	$P[A_3A_1A_2]$	$P[A_3A_2A_1]$
0.187	0.186	0.134	0.118333	0.206	0.168667
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[A_3^+]$	$P[A_3^-]$
0.373	0.287	0.252333	0.392	0.374667	0.321
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	A_3^+/mean time	A_3^-/mean time
0.377157	0.373	0.432135	0.252333	0.354847	0.374667

Table 18: $\mu_1 = 2.8, \mu_2 = 1, \mu_3 = 1.1, \sigma_{11} = 1.1, \sigma_{22} = 0.9, \sigma_{33} = 1.2$

$P[A_1A_2A_3]$	$P[A_1A_3A_2]$	$P[A_2A_1A_3]$	$P[A_2A_3A_1]$	$P[A_3A_1A_2]$	$P[A_3A_2A_1]$
0.314	0.320333	0.096667	0.027667	0.192333	0.049
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[A_3^+]$	$P[A_3^-]$
0.634333	0.076667	0.124333	0.512667	0.241333	0.410667
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	A_3^+/mean time	A_3^-/mean time
0.293931	0.634333	0.327513	0.124333	0.283072	0.241333

Table 19: $\mu_1 = 2, \mu_2 = 1.7, \mu_3 = 1.9, \sigma_{11} = 1.1, \sigma_{22} = 0.9, \sigma_{33} = 1.2$

$P[A_1A_2A_3]$	$P[A_1A_3A_2]$	$P[A_2A_1A_3]$	$P[A_2A_3A_1]$	$P[A_3A_1A_2]$	$P[A_3A_2A_1]$
0.191667	0.194	0.132	0.121667	0.205333	0.155333
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[A_3^+]$	$P[A_3^-]$
0.385667	0.277	0.253667	0.399333	0.360667	0.323667
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	A_3^+/mean time	A_3^-/mean time
0.287076	0.385667	0.319468	0.253667	0.273249	0.360667

Table 20: $\mu_1 = 1.2, \mu_2 = 1, \mu_3 = 1.1, \sigma_{11} = 3, \sigma_{22} = 2.6, \sigma_{33} = 2.5$

$P[A_1A_2A_3]$	$P[A_1A_3A_2]$	$P[A_2A_1A_3]$	$P[A_2A_3A_1]$	$P[A_3A_1A_2]$	$P[A_3A_2A_1]$
0.179667	0.195667	0.153667	0.162	0.146	0.163
$P[A_1^+]$	$P[A_1^-]$	$P[A_2^+]$	$P[A_2^-]$	$P[A_3^+]$	$P[A_3^-]$
0.375333	0.325	0.315667	0.341667	0.309	0.333333
A_1^+/mean time	A_1^-/mean time	A_2^+/mean time	A_2^-/mean time	A_3^+/mean time	A_3^-/mean time
0.108013	0.375333	0.115683	0.315667	0.12737	0.309

4 Conclusion

For each simulated model, I obtain choice probabilities for each of the options, the average reaction time of the process, the distribution of the reaction time random variable, and the distributions and average durations for each episode of the process. In the previous section, I illustrate how the choice of model parameters influences these results. We saw that increasing one of the components of the drift vector (while keeping others constant), reduces the average reaction time and increases the choice probability for the corresponding option. Moreover, we saw, that in the absence of a drift vector, increasing $||\Sigma||^2$, yields a smaller average reaction time. The simulations indicate that for the same model parameters as in Proposition 1, we get very similar results for the Disk model compared to the theoretically proven ones. I illustrate how the reaction time random variable distribution varies when only one component of model parameters changes (and when they jointly change). Next, we see how these model parameters influence the average duration of each episode. Increasing components of the drift vector yields lower average duration for each episode. Moreover, we saw that increasing only one component of the drift vector (and keeping $\mathbf{\Sigma} = \mathbf{I}$), enables the majority of simulations to end in one single episode. Lastly, I explain the main ideas and the possible scenarios/events resulting from the "Worst-best option ranking". I consider different parameters for each model and present the probabilities for each scenario and event.

5 Acknowledgements

I would like to thank my thesis supervisor, Dr Keivan Mallahi-Karai for his advice and guidance throughout the whole process. I would like to also thank Prof. Dr. Adele Diederich whose feedback on my simulations was very valuable.

A Appendix

Interactive environment for running the codes (editing possible):

1. Cube Model with two alternatives and one fall-back option

2. Disk Model

3. Cube Model with three alternatives

Non-interactive environment for presenting the outputs:

1. Cube Model with two alternatives and one fall-back option

2. Disk Model

3. Cube Model with three alternatives

References

[1] L. Koralov and Y. G. Sinai, *Theory of probability and random processes*. Springer Science & Business Media, 2007.

[2] W. H. Batchelder, H. Colonius, and E. N. Dzhafarov, *New Handbook of Mathematical Psychology: Volume 2, Modeling and Measurement*. Cambridge University Press, 2018.

[3] S. E. Alm, *Simple random walk*, http://www2.math.uu.se/~sea/kurser/stokprocmn1/slumpvandring_eng.pdf, Accessed: 2022–04-05, 2006.

[4] R. A. Wilson and F. C. Keil, "The MIT encyclopedia of the cognitive sciences," in MIT press, 2001, pp. 220–222.

[5] T. Otter, J. Johnson, J. Rieskamp, *et al.*, "Sequential sampling models of choice: Some recent advances," *Marketing letters*, vol. 19, no. 3, pp. 255–267, 2008.

[6] A. Diederich and P. Oswald, "Sequential sampling model for multiattribute choice alternatives with random attention time and processing order," *Frontiers in human neuroscience*, vol. 8, p. 697, 2014.

[7] J. A. Clithero, "Response times in economics: Looking through the lens of sequential sampling models," *Journal of Economic Psychology*, vol. 69, pp. 61–86, 2018.

[8] K. Mallahi-Karai and A. Diederich, "Decision with multiple alternatives: Geometric models in higher dimensions — the cube model," *Journal of Mathematical Psychology*, vol. 93, p. 102 294, 2019, ISSN: 0022-2496. DOI: https://doi.org/10.1016/j.jmp.2019.102294. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S002224961930152X.

[9] ——, "Decision with multiple alternatives: Geometric models in higher dimensions — the disk model," *Journal of Mathematical Psychology*, vol. 100, pp. 102–493, 2021, ISSN: 0022-2496. DOI: https://doi.org/10.1016/j.jmp.2020.102493. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022249620301164.

[10] A. Tversky, "Elimination by aspects: A theory of choice.," *Psychological review*, vol. 79, no. 4, p. 281, 1972.

[11] J. Murphy, *Simulating brownian motion in python*, https://www.youtube.com/watch?v=hqSnruUe3tA, Accessed: 2022–04-10, 2020.

[12] *Seaborn.distplot*, https://seaborn.pydata.org/generated/seaborn.distplot.html, Accessed: 2022-04-16.

[13] *Seaborn.jointplot*, https://seaborn.pydata.org/generated/seaborn.jointplot.html, Accessed: 2022-04-16.

[14] *3d surface (colormap)*, https://matplotlib.org/stable/gallery/mplot3d/surface3d.html, Accessed: 2022-04-21.