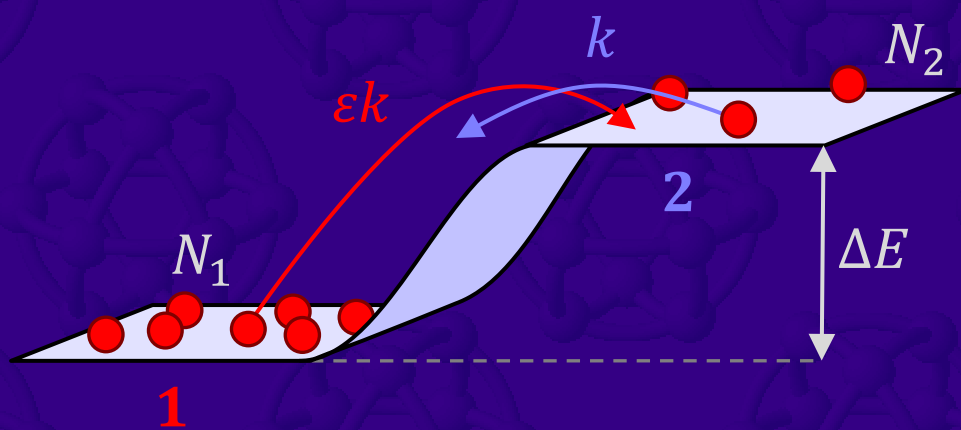
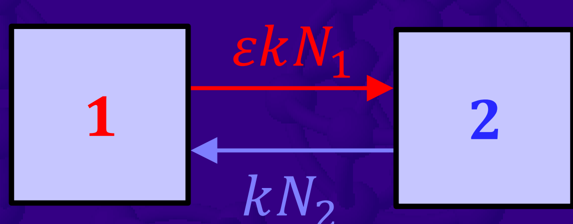
Biophysics and Physiological Modeling

Discovering Science with the Marble Game





Peter Hugo Nelson

This material is based upon work supported by the National Science Foundation under Grant Nos. 0836833, 1817282, and 2306506. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

Note for students – using this book

Welcome everybody. The materials for this book were developed as part of a National Science Foundation (NSF) sponsored program to transform undergraduate science education from a sometimes-boring passive lecture-based process into an active-learning journey that makes physics relevant to the life-sciences. When you work through the chapters of this book, you'll gain new insights into important topics in biochemistry, biophysics, molecular biology, physiology, and physics.

Reading is not enough! The text includes *interactive* spreadsheet activities that are essential to your learning. You should have a computer close by so that you can do those activities in Microsoft Excel® or a compatible program. That way you'll be able to *discover* for yourself what the spreadsheet models actually do. I.e., to get the most out of this book, you should do the activities and answer the questions as you go along. The questions are an integral part of the **active learning process** that's designed to help you gain insights into life at the molecular level. On the companion website, I've posted videos to help you get started. While most parts look like a regular science textbook (or sometimes a computer guidebook), the main purpose is to give you the basic information that you need to do the activities and discover what they tell us about how the molecules of life work. You can think of the chapters as being "Excel labs". If you read a lab manual without actually doing the lab, you'll get some idea of what's going on, but most people learn better by *just doing it*!

After your spreadsheet is working, you'll be asked follow-up questions that are intended to make you think about what's actually happening in your spreadsheet and what it means. Sometimes those questions can be very straightforward – their purpose could be as simple as making sure that you really did look at (and think about) the spreadsheet thing that you just changed. Other times, you'll have to think carefully about what your spreadsheet is doing to formulate your own answer. The goal is to guide you towards a better understanding of what's actually happening in your spreadsheet and what it means. That way, you'll *discover* for yourself some of the foundational scientific principles used in biochemistry, molecular biology, physiology and physics. **DISCUSSION QUESTIONS** may be discussed in class. **CALCULUS QUESTIONS** are optional for algebra-based courses. **RESEARCH** and **CHALLENGE QUESTIONS** are extra credit. The questions are sometimes very straightforward – their purpose can be as simple as making sure that you really did look at (and think about) the spreadsheet thing that you just changed.

Often there'll be follow-up information that will only make sense after you've done the activity. This additional information is contained in **About what you discovered** (**AWYD**) sections. Please read this information *after* you've attempted the preceding questions.

About what you discovered: chapter appendix information and graphs

If you get stuck on a question or in a spreadsheet, read ahead into the "About what you discovered" (AWYD) section – the information you find there might be helpful! However, the primary purpose of the AWYD sections is to further discuss the meaning of the activity you just completed. This material will usually make much more sense *after* you've completed the activity and answered the follow-up questions.

In the AWYD section, there may be a hypertext link to an answer graph in the chapter appendix. After the graph, there'll be a hypertext link back to the main text of the chapter. If you have this PDF open, you can click the hyperlink. If you're reading hardcopy, you can follow the hyperlinks by flipping to the appendix at the end of the chapter.

This book uses a number of different fonts to indicate different type of information, e.g. **glossary terms**, web links, **Excel** terms, **Q.1.1** (questions), *action words*, references to **SECTION HEADINGS** and equations like $N = N_1 + N_2$. If you run across a font in the book and can't remember what it signifies check out **APPENDIX A.1** or come back here and review this **NOTE FOR STUDENTS**. If you come across a symbol like N_1 and can't remember what it stands for, you can look it up in the **GLOSSARY OF SYMBOLS** in the **APPENDICES** at the end of the book. The appendices also contain other useful information such as a list of units, physical constants, and unit conversions etc. ❑

Biophysics and Physiological Modeling
Chapter 1: Introduction – the marble game

v.4.8 © pHn 2024

Introduction: a breath of fresh air

In this book, we're going to spend a lot of time thinking about how the molecules of life behave. There's usually a lot going on, so we'll often focus on a single molecule. Imagine…

> *Your little sister Bonnie has just received the long-awaited animated movie –* ***OXYGEN STORY 2****. It's a reality-based cartoon that follows the continuing adventures of Dion – Bonnie's favorite oxygen molecule … While on vacation sailing around the atmosphere, Dion wanders near a friendly looking human. He gets caught up in a crowd and is pushed in. After a wild and turbulent ride, Dion finds himself in a hot and steamy cave. The walls are wet and soapy. After bouncing around in the cave for a while, Dion discovers that he can slip right through the cave wall into a salty red river…*

While watching the movie with Bonnie, you realize that Dion's passage through the wall represents an interesting (and important) example of molecular transport across a boundary in physiology. The steamy cave is one of the small air sacs (alveoli) in the human's lungs and the red river is blood plasma in a small blood vessel (capillary) right next to the air sac. Fig.1.1 shows a simplified representation of the oxygen (O_2) molecule's transition from the air into the blood plasma. Box 1 is a tiny portion of the blood plasma and box 2 is an equally tiny portion of the breath gases in the air sac. The gap between the boxes represents the tissue separating the air from the plasma. This tissue is **permeable** to O_2, i.e., the O_2 molecule can pass through.

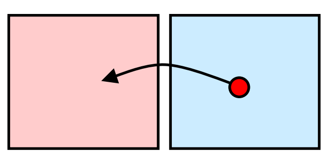


Fig.1.1 Schematic representation of an O_2 molecule moving from the air in your lungs to your blood plasma.

When the human inhales, fresh air rushes into box 2. Once the air is in box 2, the process is entirely automatic. At any given moment, the O_2 molecule has a random chance of jumping from the air (box 2) into the blood plasma (box 1) so that it can start its long, life-giving journey…

This description of how O_2 gets into the bloodstream seems straight forward, but there's a problem. The jump that moves the O_2 molecule from the air into the plasma is entirely **passive**, implying that the reverse jump can also happen spontaneously as shown in Fig.1.2. This reversibility shouldn't affect our argument – so long as the inward jump (Fig.1.1) is more likely than the reverse outward jump (Fig.1.2) …but there's a catch… a *big* one! The outward jump in Fig.1.2 is actually *40 times more likely* than the inward jump in Fig.1.1! This means that once the O_2 molecule is in the plasma box, it's 40 times more likely to go back into the air than it was to enter the blood plasma in the first place – but how can that be? The human needs that O_2 to stay in the blood so that it can be delivered to where it's needed.

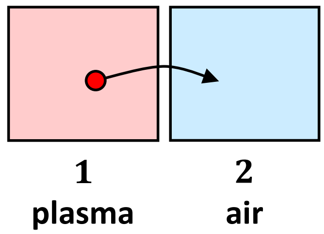


Fig.1.2 Schematic representation of the reverse process – an O_2 molecule moving from your blood plasma back into the air in your lungs.

Physiology is full of descriptions of molecules moving around and functioning in ways that are essential for us to stay alive. Many are reversible processes similar to the jumps discussed above. This book will help us to discover what's actually going on – at the level of individual molecules. This type of description is not limited to physiology; it's also applicable to most of science and engineering.

The marble game

Fig.1.3 Photo of the marble game. It has ten marbles that can jump between two boxes. A ten-sided dice is rolled to pick which marble jumps next.

The main goal of this book is for us to learn *how to think* about these problems. Along the way, we'll *discover* how the apparently contradictory statements made above can be combined into a coherent explanation. But we'll start by learning how to play a children's marble game invented by the author – see Fig.1.3 [Nelson 2012]. It's a rather simple game with marbles jumping between two boxes, but don't be fooled by just how simple it is! Once we've discovered how it works, we'll be able to modify it (in **CHAPTER 3**) to discover how Dion's jump through the wall really is consistent with O_2 uptake in humans, even though the reverse jump really is 40 times more likely.

The purpose of this **CHAPTER 1** is to introduce you to the marble game and to using a spreadsheet (like Microsoft Excel®). The chapter is a "how-to" guide for Excel. It's assumed that you already know how to open Excel and create a simple scientific graph (Excel calls them charts). There are videos posted online to help get you up to speed. In **SECTION 1.1** we'll start by learning how to play the marble game – *by hand*. **SECTION 1.2** introduces Excel and the function **RANDBETWEEN** that can be used to simulate the rolling of the ten-sided dice[1] used in the game. In **SECTION 1.3** we'll teach Excel how to play the game automatically for thousands of turns. That way, we'll be able to analyze what happens after thousands of turns and with hundreds of marbles without wasting lots of time (and without getting bored silly) doing it by hand! In **SECTION 1.4** we'll discover that the marble game can explain the fundamentals of "dynamic equilibrium" in biochemistry and physiology. In **SECTION 1.5** we'll discover that the *unbiased* jumps in the marble game produce a net transfer from high to low concentration. This observation is known as "Fick's law of diffusion". We'll then briefly discuss what the marble game can tell us about vital physiological processes including the distribution of O_2, CO_2 and glucose. What we discover in this chapter will help us understand everything that follows – yes, it's really that important! In later chapters, we'll see how the marble game can be modified to investigate a wide variety of other important physiological processes including:

- membrane transport
- distribution of O_2, CO_2 and glucose
- drug delivery & elimination
- osmosis and osmotic pressure
- bioelectricity & membrane potentials
- diffusion and random walks
- molecular signaling & calcium
- synapses & neurotransmitters
- ion channel permeation
- ion channel gating & action potential

In **CHAPTER 12** we'll apply the marble game to epidemiology – specifically the spread of the COVID-19 pandemic in 2020. In **CHAPTER 13** we'll discover how the same computational methods can be used to model (and hence understand) the foundational concepts of traditional physics and engineering. We'll model the performance of the world's fastest production car.

Physical basis of the marble game

Before we start playing the game, let's talk a little about the physical basis for the **jump** of a marble from one box to the other shown in Fig.1.4. The marble represents one of the molecules that we're interested in, and the boxes represent two tiny regions of space. In the simplest case, the two boxes could represent two adjacent regions of intracellular fluid (cytosol), or they could represent two different regions separated by a membrane (e.g., interstitial fluid and cytosol). Because of random thermal motion, all the molecules in the two boxes are in constant motion. The net result is a **passive process** that results in a particular molecule jiggling *randomly* through the solution in a process that's unpredictable in advance like rolling dice. This **Brownian motion**

[1] According to the Cambridge Dictionary and Oxford Dictionary, "dice" is both singular and plural.

is the physical mechanism for the jumps between the boxes in the marble game. As we'll discuss in more detail in **CHAPTER 10**, the random motion produced by the marble game **simulates** this Brownian motion. This random motion produces "diffusion", which we'll be studying in more detail in **CHAPTERS 7 AND 10**. Since all biological molecules and ions move in a comparable manner, it's really important for us to understand how this random motion works in as much detail as possible!

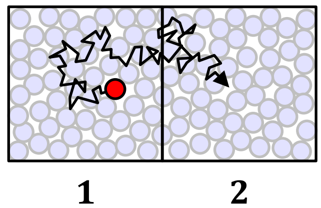


Fig.1.4 Schematic representation of the physical basis for the marble game. Because of random thermal motion, all the molecules are jiggled around constantly. If we focus on a single tagged molecule, this random jiggling produces **Brownian motion**. The black line is a "snail trail" of one possible trajectory that transfers the molecule from box 1 → 2. The net effect of this trajectory is that the molecule **jumps** from box 1 → 2 during a short period of time.

About what you discovered: Brownian motion demo

On the companion website (https://circle4.com/biophysics/) to this book there's a link to the Web VPython program MarbleGame [Nelson 2020]. It's an animated **Brownian Motion simulation** (**BM sim**) that you can run in the web browser of your favorite device (PC, Mac, phone, tablet etc.) It animates the physical basis of the marble game with $N = 50$ tagged marbles (representing O_2 molecules) executing **Brownian motion** because of collisions with the water molecules surrounding them (not shown). The marbles don't interact with each other at all. Each marble randomly walks around both boxes – as if the other marbles aren't even there – marble independence! This program combines ideas from **CHAPTERS 1**, **2**, **3** and **10**.

On the companion website there's also an Excel file BPM.Ch01_BrownianMotionDemo.xlsx that includes an interactive Excel simulation of Brownian motion. *Open it now*. The simulation shows a **snail trail graph** of a 2000-step random walk (**CHAPTER 10**). The two boxes are physically identical and the dotted line between them has no physical significance. It's just an imaginary dividing line. The green circle is the starting point, and the red circle is the ending point (2000 steps later). The snail trail represents the random Brownian motion of a single molecule such as O_2 in blood plasma, cytosol, or intracellular fluid etc. This random motion is the physical basis for the jumps in the marble game and for the "diffusion" of oxygen as we'll discover later in this chapter. If you open the file in Excel, pressing the F9 function key will make a new random walk starting out where you left off. Pressing F9 repeatedly will give you an idea of how Brownian motion produces random walks that can periodically make the molecule "jump" between boxes. **Warning:** You should close the Excel program completely when you're finished with the Excel Demo. If you don't, the special features used in the spreadsheet will mess up what follows. ❑

About what you discovered: diffusion *vs.* convection

It's common for people to confuse *diffusion* with *convection*. Diffusion is a process that is hidden way down at the molecular level and requires us to consider the motion of individual molecules as we'll do with the marble game. However, convection is literally the breath of life. Hold the palm of your hand a few inches in front of your face and blow on it gently. The muscles in your diaphragm contract and compress the air in your lungs. This physical force slightly increases the concentration of air molecules in your lungs, but the flow of air that you feel on your palm is *not* the result of diffusion – it's **convection**. The contractile force of your diaphragm increases the air pressure in the lungs, which literally pushes the air out of your lungs. The air molecules flowing up through your trachea and then out through your mouth are all being pushed in the *same direction* by the pressure difference – producing a **convective flow**. The air molecules move together just like the water flowing out of a tap, or the blood flowing in your veins. Diffusion is not like that. The forces producing Brownian motion are completely random and have no average direction. Unlike convection, the diffusive forces on a particular molecule are completely **uncorrelated** so that the forces on one molecule produced by collisions with adjacent molecules average out to zero in the long run. ❑

1.1 Playing the game

The marble game is a single-player game that has two boxes labeled 1 and 2 and ten marbles that can either be in box 1 or box 2. Fig.1.5 is a **schematic representation** of the marble game. At every turn, we'll roll a ten-sided dice (see Fig.1.3) to randomly select one of the ten marbles to jump to the other box. This simulates random Brownian motion as each marble always has an equal probability of jumping to the other box. At the beginning of every turn, we'll assign a number to each of the ten marbles. The marbles in box 1 are labeled 1 through 3 and the marbles in box 2 are labeled 4 through 10. The purpose of the marble game is to keep track of N_1 – the number of marbles in box 1. $N_1 = 3$ summarizes the arrangement of marbles shown in Fig.1.5 because there are 3 marbles in box 1 (and hence 7 marbles in box 2).

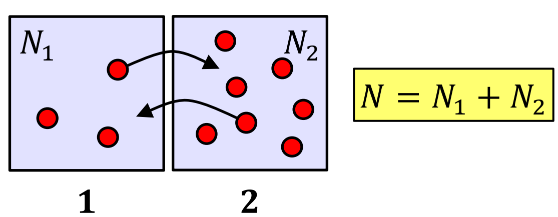


Fig.1.5 Schematic representation of the marble game. N_1 is the number of marbles in box 1, N_2 is the number of marbles in box 2, and there is a total of $N = N_1 + N_2$ marbles in the game. The figure shows an $N = 10$ marble game with $N_1 = 3$ marbles in box 1 and $N_2 = 7$ marbles in box 2.

In the marble game, you roll a ten-sided dice to randomly pick one of the numbered marbles. It then "jumps" to the other box. Let's call the random number that you rolled r. This random number r determines which marble jumps next. For the arrangement shown in Fig.1.5, if we roll a number between 1 and 3 (inclusive) (i.e., $r \leq N_1$), then a marble in box 1 is selected. If we roll a number between 4 through 10 (inclusive) (i.e., $r \nleq N_1$), then a marble in box 2 is selected. These are the only two choices. Hence, if $r \leq N_1$, a marble jumps from box $1 \rightarrow 2$. Otherwise, a marble jumps from box $2 \rightarrow 1$. At first glance, the marble game rule might seem like its biased, but it's not. Each one of the numbers r rolled on the dice are equally likely, so that each of the marbles has a one-in-ten chance of being selected to jump next. This realistically simulates ten **independent** Brownian particles that each move completely independently of the others because their individual chance of jumping next is always one in ten – independent of how the marbles are distributed between the two boxes.

About what you discovered: translating math into English

Whenever you see a mathematical symbol in the text of this book, a light should go off above your head (well at least figuratively!). Mathematical symbols like N_1 and r are used because they are convenient shorthand for us to use to communicate with each other efficiently (just like the text messaging shorthand "lol"). When you see "N_1", you should think "the number of marbles in box 1". When you see "r", you should think "the random number rolled by the dice". (Please note that the underlining in the previous two sentences is intended to help you connect the symbols with the things that they represent.) When these symbols are used in a mathematical expression, the expression is shorthand for a whole sentence – that's *always important*. E.g., "$r \nleq N_1$" should be pronounced "the random number rolled by the dice is *not less than or equal to* the number of marbles in box 1". If you didn't actually do the translation from math shorthand into English when you read the previous paragraph in the text, then you should go back and reread it!

Fig.1.5 includes an equation

$$N = N_1 + N_2 \tag{1.1}$$

It's a relationship between three quantities that's *always* true. We've already talked about N_1, it's the number of marbles in box 1. N_2 is the number of marbles in box 2, and N (with no subscript) is the **total number of marbles** in the game. Equation (1.1) can be translated into English as "the total number of marbles in the game is the same as the sum of the numbers in box 1 and box 2". It's important that you actually do the translation into English when you read the equation – at least until you become fluent in the language of math. We'll call equation (1.1) the **bookkeeping equation** because it reflects that the fact that the number of marbles in boxes 1 and 2 must add up to the total number of marbles N. ❑

As a concrete example, let's consider an $N = 10$ marble game that starts out with $N_1 = 3$ marbles in box 1. At the beginning of the game, you place $N_1 = 3$ marbles in box 1 and $N_2 = 7$ marbles in box 2. The purpose of the game is to make a table for N_1 – the number of marbles in box 1 as N_1 changes at every turn (see Table 1.1). The marble game has just one repeated rule:

Marble game rule

- Roll the ten-sided dice. If r is *less than or equal to* N_1 ($r \leq N_1$), then move a marble from box $1 \rightarrow 2$, otherwise move a marble from box $2 \rightarrow 1$.

That's it! Table 1.1 shows the result of a **sample** game after two turns. Let's see what it means…

Table 1.1 Record of the progress of a sample marble game.

Turn	r	N_1
0		3
1	6	4
2	4	3

In turn 0 we put three marbles in box 1 and seven marbles in box 2 so that $N_1 = 3$. In turn 1, we rolled $r = 6$ on our ten-sided dice. We then compare this with N_1 (the current number of marbles in box 1) to determine whether we choose a marble in box 1 or box 2 to jump next. In this case $r \nleq N_1$ (i.e., $6 \nleq 3$), which means the chosen marble (number 6) in is box 2 and we move a marble from box 2 → 1. This gives $N_1 = 4$ (at the end of turn 1) as we added a marble to box 1. In the second turn, $r = 4$ is rolled. In this case $r \leq N_1$ (i.e., $4 \leq 4$) so that the chosen marble is in box 1 and we move a marble from box 1 → 2, resulting in $N_1 = 3$ at the end of turn 2.

Before we start playing the game in earnest, let's talk a little more about what it means. When we roll the dice, we use the number on the dice to randomly pick one of the ten marbles and then move it to the other box. Hence, at every turn, all the marbles are treated in exactly the same way, and they all have the same probability of moving to the other box (one in ten). The probability is the same for jumps from box 1 → 2 as from box 2 → 1. The physical implication is that the molecular environments are similar in the two boxes, e.g., for an O_2 molecule the boxes could be blood plasma, interstitial fluid or cytosol (all of which are similar salty solutions from Dion's perspective).

Q.1.1 A fresh game is started. In turn 0, three marbles are placed in box 1. A ten-sided dice is rolled, and the following sequence of numbers comes up r = 6, 4, 4, 7, 6, … Using the rules of the game, *write out by hand* a table of N_1 values generated in this sample game for turns 0 through 5.
Note: If you have an electronic writing device, you can cut and paste the ink into MS Word. Otherwise, take a picture of your pencil-and-paper answer with a smartphone and paste it into Word. Then in Word's **Help**, search on **compress pictures** to reduce the picture resolution so that your Word file doesn't get huge. Don't forget to label your answer with the question number.
Hint: While you're filling out your table you should carefully *check* that you get each entry in Table 1.1 and make sure that you *know why* each entry has the value shown. You should also *note* that no random number r is needed for turn 0, as we haven't made a move yet.

Q.1.2 Another new game is started. Once again there are three marbles in box 1 in turn 0. A ten-sided dice is rolled, and the following sequence of numbers comes up r = 3, 2, 1, 1, 10, … Using the rules of the game, *write out* a table of N_1 values generated in this sample game for turns 0 through 5.

Q.1.3 *Sketch by hand* a graph showing N_1 versus turn for the game of Q.1.1 and for the game of Q.1.2 (N_1 on the y-axis and turn on the x-axis). Plot both games on a single graph. Show the values of N_1 as symbols (or markers) (e.g., filled circles and squares) and connect them together with dashed straight lines (as a guide to the eye). Be sure to *label* your graph correctly using scientific format and then *paste* your answer into Word using the technique described in Q.1.1.
Hint: The hand-drawn graph should include a title, axis labels (with units if applicable) and the axis numbers should have axis tick marks. Your graph should also include a **legend** to indicate which line corresponds to each game.
Note: Strictly speaking, the graph should not have lines "connecting the dots". These are added here just as a guide to the eye – to help you see what happened in each of the two games as a function of time.

About what you discovered: losing all your marbles!

What you should have discovered in Q.1.2 is that it's possible to end up with no marbles in box 1 (i.e., $N_1 = 0$). It's important to note that the rules of the game work just fine in this case. When we roll the dice, we randomly pick one of the ten marbles and then move it to the other box, but because all the marbles are in box 2, the only move that's possible is for a marble to move from box 2 → 1 and we get one of the marbles back in box 1. ❑

We'll want to play the game for a thousand turns or so and to do this over and over to see what happens. Rather than getting bored silly trying to do this by hand … we'll teach Excel to play the game for us and to keep track of the game history so that we can discover the consequences of the game rules – for games that are much longer than we would ever want to play by hand.

Why Excel?

In this chapter we'll discuss using Microsoft Excel®, but you can use another compatible spreadsheet program if you like. There are many other (more powerful) choices such as MATLAB, or Maple, or programming languages such as Visual Basic, C++, Java, or Python, but none of these are as widely used, or are as easy to learn, as Excel. People in all walks of life use Excel to perform all kinds of routine tasks ranging from generating a list of clothes to take on a business trip (yes, my wife actually does this!) to the budget of the United States Government (see www.cbo.gov). It's also easy to generate graphs of data within Excel, and we'll use this feature to discover what all those numbers in the spreadsheet actually mean. Excel is sort of like a big Swiss Army knife. It may not be the best tool for every job (and it might not even be a very good knife), but you can use it for a lot of tasks…

Getting started using Excel

If you still have the Brownian motion Excel demo spreadsheet open, please *close* it now and then *close* the Excel program.[2] Then open the Excel app and then *open* a **Blank workbook**. In the top left-hand corner of the (blank) spreadsheet, *type* the heading "**Turn**" in **cell A1**. Cell **A1** is the empty box located in **row 1** of **column A** of the spreadsheet. After you *press* the ENTER key, your spreadsheet should look something like Fig.1.6, depending on the version of Excel you're using.

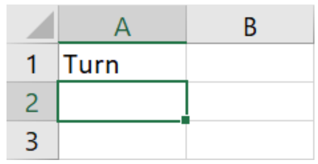


Fig.1.6 Screenshot from Excel.

About what you discovered: different versions of Excel

When I was writing this chapter, I was using Excel 2016. All the figures and instructions relate to the version I was using. Other versions of Excel will look a little different and unfortunately, the exact instructions for how to do things – like adding a chart – probably won't be exactly the same, but hopefully you'll get the idea… (BTW you might have to change the **View** to **Normal** to see something similar to Fig.1.6 – see Excel **Help**.) ❑

Next, *enter* the heading ***r*** (for random number) in cell **B1** and the heading N_1 in cell **C1**. It's okay if you don't want to bother with the subscript formatting. You can just type **N_1** instead; the formatting makes no difference to how the spreadsheet works, but it does make it look more scientific. Now *enter* the number **0** in cell **A2** and **3** in cell **C2**. We don't need a number for r yet, as we are still setting up and haven't started turn 1 yet.

Now *enter* **=A2+1** in cell **A3**, but don't hit the ENTER key just yet! Your spreadsheet should now look something like the first screenshot in Fig.1.7 *before* you hit ENTER, and like the second screenshot *after* you *hit* the ENTER key.

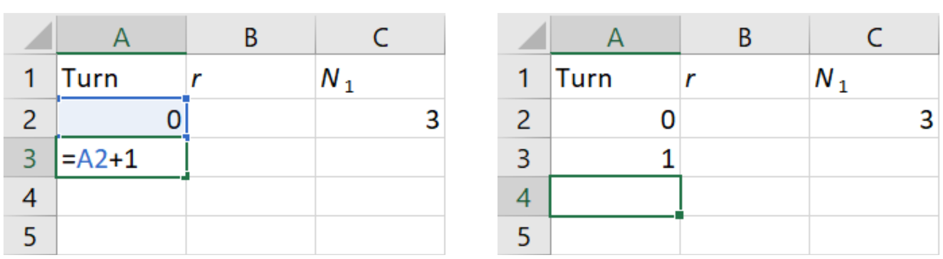


Fig.1.7 Screenshots from Excel, before and after hitting ENTER (see text).

The number **0** in cell **A2** is just a regular number, but the **formula =A2+1** describes how the value of cell **A3** is to be calculated. In words, the recipe reads, "take the value of the cell above (**A2** in this case) and add the **number 1** to it". When you press the ENTER key, Excel does the calculation

[2] The special formatting in the Brownian motion demo spreadsheet may mess up your work. Closing and reopening the Excel application, should avoid that potential problem.

and displays the calculated **value** instead of the formula (see the after figure on the right of Fig.1.7). Excel recognizes that cell **A3** contains a formula because it starts with an equals sign **=**. If you forget the **=** at the beginning, Excel will think that you are just typing *words* instead of a *formula* to be calculated. The result would then be that Excel would just display what you typed – without doing anything with it. The moral of the story is simple … don't forget the **=** at the beginning of formulas.

We now want to add a formula for the random number *r* in column **B**. Excel has a function **RANDBETWEEN** that produces a **random number** exactly like rolling the dice. *Add* this function to the spreadsheet by *typing* **=randbetween(1,10)** in cell **B3** (don't forget the **=**). After you do that, your spreadsheet should look something like the first screenshot in Fig.1.8 *before* you hit enter, and like the second screenshot *after* you hit the ENTER key. Don't worry if the number you get in cell **B2** is different from the "**5**" in Fig.1.8, that's perfectly normal – it is a random number after all. However, it should be an integer between 1 and 10.

	A	B	C	D
1	Turn	*r*	N_1	
2	0		3	
3	1	=randbetween(1,10)		
4				
5				

	A	B	C	D
1	Turn	*r*	N_1	
2	0		3	
3	1	5		
4				
5				

Fig.1.8 Screenshots from Excel, before and after hitting ENTER (see text).

About what you discovered: loading RANDBETWEEN into Excel

If you get an error message like **#VALUE!** or **#NAME?** when you type **=RANDBETWEEN(1,10)** in a blank cell, that means Excel did not understand what you typed. First carefully double check what you typed is the exact same sequence of characters and then hit the ENTER key. If you still get the error message, the most likely cause is that the (free) **Analysis ToolPak** is not installed. To get instructions on what you need to do, go into **Excel Help** and search on **RANDBETWEEN**. You should find instructions on how to load the **Analysis ToolPak** (sometimes called **Data Analysis Tool**, or **Data Analysis ToolPak**). If none of that works and you still can't get the **Analysis ToolPak** to load, you can use **=INT(10*RAND()+1)** instead of **RANDBETWEEN(1,10)**. You can use Excel's **Help** feature to figure out what the functions **INT** and **RAND** do. ❑

The formula **=randbetween(1,10)** determines how the value of cell **B3** is to be calculated. In words, the recipe reads "make the value of the cell equal to the value of the function **RANDBETWEEN(1,10)**". When you press the ENTER key, Excel does the calculation and displays the calculated value instead of the formula (see the right panel in Fig.1.8). Excel has a large number of built-in functions like **RANDBETWEEN**. Some need one **argument** like the trigonometric functions e.g. **SIN(3.14159265)**, where the angle **3.14159265** (in radians) is the argument of the function. The function **RANDBETWEEN** is a function that needs two arguments. Let's discover what it does…

1.2 RANDBETWEEN(1,10) does what?

For the moment, we'll forget about the marble game and focus on **discovering** what *r* **=RANDBETWEEN(1,10)** does – by investigating how it works in an Excel spreadsheet. As a heads up, we'll be following this same **discovery-based approach** over and over again in the rest of the book.

> **Q.1.4** Leaving the cursor where it is, *hit* the DELETE key in the empty cell **B4**., then *briefly describe* (in words) what happens in cell **B3** when you hit DELETE.

> **Q.1.5** *Describe* what happens if you hold down the DELETE key? (or press it repeatedly?)

After you've answered Q.1.5, check out:

About what you discovered: DELETE in a blank cell (F9)

When you hit the DELETE key in a blank cell, you trick Excel into thinking that you changed that cell in the spreadsheet. Hence, Excel recalculates all the cells in the spreadsheet and revaluates all the functions including the **RANDBETWEEN(1,10)** function. That's equivalent to rerolling our ten-sided dice. We'll be using this technique often. (BTW you can press the F9 function key and get the same result.) These instructions are for an IBM-compatible PC. If you have an Apple (or other) device, you'll have to find the equivalent procedure in Excel's **Help**. On the iPad there is a **Recalculate** button under the **Formulas** tab. ❑

Highlight cells **A3** and **B3**. There are lots of ways of doing that: one way is to *left click-drag* the mouse over the center of cells **A3** and **B3**; another way is to *highlight* cell **A3** and *press* shift + RIGHT ARROW once. (For more information see the Excel's **Help** feature.) Your spreadsheet should now look something like Fig.1.9.

	A	B	C
1	Turn	*r*	N_1
2	0		3
3	1	5	
4			

Fig.1.9 Screenshot from Excel.

Hover the cursor over the bottom right-hand corner of the highlighted area. The mouse pointer should change from the normal open cross ✢ to a smaller solid cross +. Then *left-click-drag* down two rows to row **5**. When you *release* the mouse, your spreadsheet should look something like Fig.1.10. Once again, the numbers you'll get for r will be different from the figure. If you select an empty cell (such as **A6**) and press DELETE, then the r values should change (see the second screenshot in Fig.1.10.).

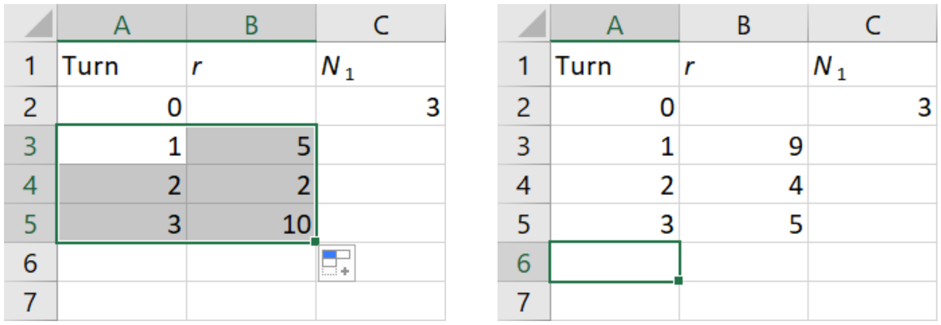


Fig.1.10 Screenshots from Excel, before and after hitting DELETE (see text).

So why did we do this in such a complicated way? …We could have just typed 0, 1, 2, 3 in column **A** and **=randbetween(1,10)** in the cells of column **B**. The answer is that if we write the spreadsheet this way, we can **Copy** all of row **2** – any number of times we like – without any extra effort! So how does that copy work? …To answer that question, we'll need to talk about a bar – the **Formula bar**. It's a place where you can go to get information about functions and the content of cells in the spreadsheet.

Formula Bar

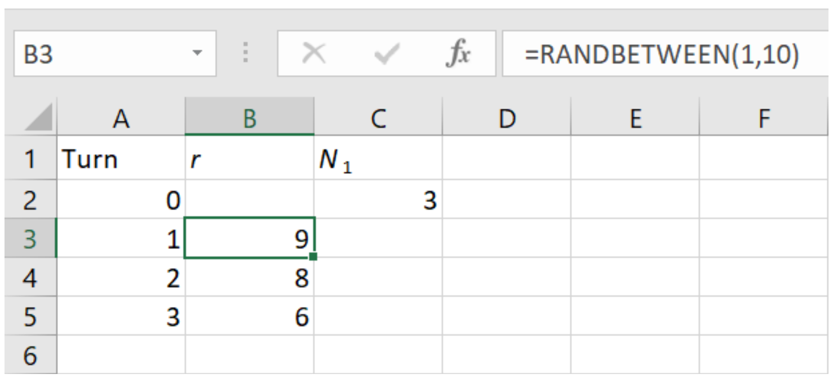


Fig.1.11 Screenshot from Excel.

When you enter a formula into a spreadsheet cell, what you type is also repeated in the **Formula Bar** (usually located above row **1** of the spreadsheet). *Click* in cell **B3** and you should see the formula shown in Fig.1.11 in the **Formula bar**. *Notice* that the formula bar displays the formula that we originally typed into cell **B3**, although Excel has capitalized the **RANDBETWEEN(1,10)**, because Excel has recognized it as a built-in function. Once a formula has been entered in a cell, you can go to the **Formula bar** to edit the existing formula. You can also get information about how any cell is calculated. For example, if you *select* **A3** and then *left click* inside the **Formula bar,** then you'll get something like Fig.1.12.

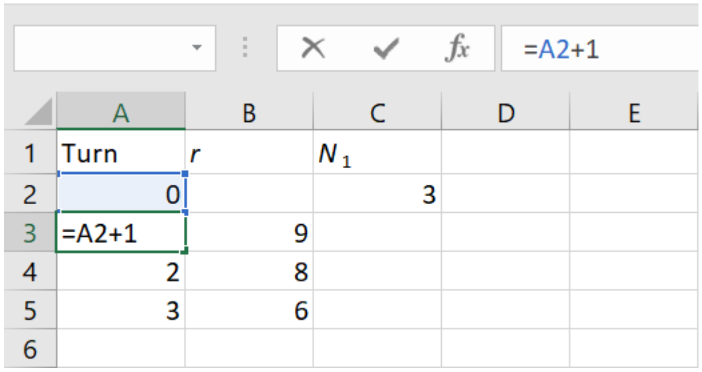


Fig.1.12 Screenshot from Excel.

When you *left click* in the **Formula bar**, Excel highlights the cells that are being used in the calculation. When *looking at* cell **A3**, the formula is **=A2+1** and cell **A2** is highlighted in blue in both the spreadsheet (blue box) and in the **Formula bar** (blue font for **A2**). This feature is very useful for checking what the spreadsheet is actually doing. Unfortunately, the main reason you'll want to do that is because your spreadsheet isn't working correctly. The best way to check a spreadsheet is to go through each cell in turn and make sure that it's really doing what you intended.

Okay, so when we copied row **3**, what actually happened? If you *select* **A5** and *click* inside the **Formula bar**, you'll get something like Fig.1.13. This figure shows the formula used in cell **A5** i.e., **=A4+1**. In words, the recipe reads "take the value of the cell above (**A4** in this case) and add the number **1** to it". The recipe – in words – is *exactly the same* as the recipe for cell **A2**. What Excel did when we copied row **3**, was to copy the *formula* and not the *value* (the number **1** in this case). Relative to cell **A5**, the cell used to start the formula is the one directly above it (namely **A4**). During the copy, the relative relationship between the cells is maintained. We'll call this way of finding one cell relative to another – **relative cell referencing**. This is the standard way that Excel interprets formulas during a copy.

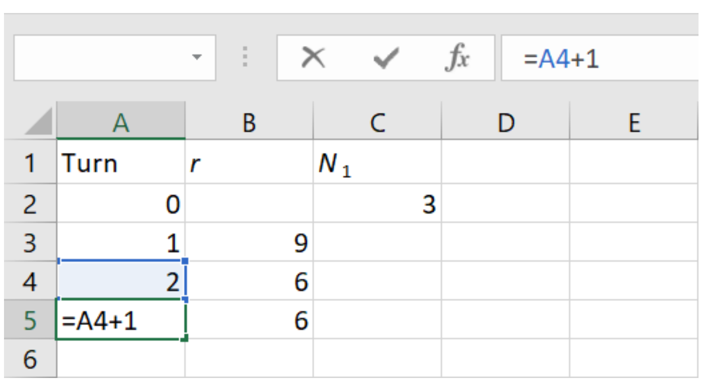


Fig.1.13 Screenshot from Excel.

Because the formula in column **B** doesn't refer to any other cells, the formula doesn't change when it's copied. You can *confirm* that by *clicking* on cell **B5** in your spreadsheet.

Q.1.6 *Write out* the exact formula in each of cells **A4** and **B4** in your spreadsheet.
Hint: use the **Formula bar** to see them.

If all has gone well so far, now would be a good time for you to *save* a copy of your spreadsheet in a folder for this chapter as a **backup**. It doesn't really matter what you call the spreadsheet, but something obvious like "BPM.Q.1.6" would be a good idea. I added the ".Q.1.6" to indicate that I saved the spreadsheet after completing Q.1.6. I strongly recommend that you save early and often, and to *different* file names. It's very easy to do something inadvertently that ruins the whole spreadsheet. If that happens and you can't **Undo**, then you'll have to go back to a previously saved version of the spreadsheet. At this point, you should also *save* a new working copy of your spreadsheet as "BPM.Q.1.7" so that you don't over-write your "BPM.Q.1.6" backup.

Now, *extend* your spreadsheet to 100 turns by *copying* all of row **5** (cells **A5** and **B5**). *Plot* a graph of the r column versus turn number (this means put r on the y-axis and turn on the x-axis) on the same **sheet** as your calculations. When you're done, your graph should look something like Fig.1.14. Once again, the random numbers you get for r will be different from the figure.

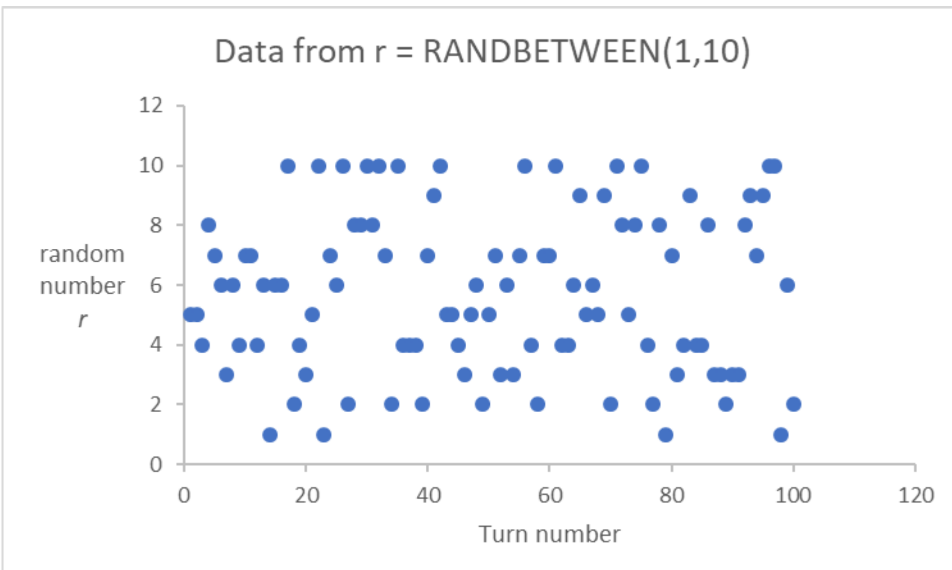


Fig.1.14 Excel **X Y (Scatter)** chart of the output of the function **RANDBETWEEN(1,10)**. The **Data Series** has been **Format**ted as **Scatter with only Markers**.

There's more than one way to make the graph and unfortunately the methods depend on the version of Excel you are using. The following methods worked for me… *Highlight* all the cells corresponding to turn and r for turns $1 - 100$ (*using left click-drag*), then

[Insert] > [Charts] > [Scatter]

…or you can try…

[Insert] > [Recommended Charts] > [Scatter] > [Scatter]

You should use a **Scatter Chart**. **Note:** The **Line or Area** type of **Chart** is *not* what we want, as we want to specify both **X** and **Y** from our spreadsheet. With the chart *on the same sheet*, you'll be able to see the chart change as you update anything in the spreadsheet. Instant gratification!

Before you go any further, you should *double check* that there is only one **series** plotted in your chart (if you didn't highlight the correct data before making the chart, you'll probably get two series – one for r and one for N_1). You can *check* by *right clicking* on the chart and then *using* **Select Data…** You should also *check* that the x-axis shows turn numbers running from $1 - 100$, and that the y-axis shows random numbers r between 1 and 10.

Select the **Chart Tools > Design** tab, and then *use* **Add Chart Element** to *enter* **axis titles** and a **chart title** that are clear, descriptive, and have the correct units (in this case there are no units but remember a graph with missing units – is just… *wrong*!). If you have trouble remembering how to do that – or figuring out how to do it in a new version of Excel – you should use Excel's **Help** feature and search on "chart title" or "axis title". *Adjust* the Excel graph to match standard scientific practice by *removing* any gridlines etc... You can *remove* the gridlines by *highlighting* them and then *deleting* them. You'll need to *select* the axis to *format* it with the style of tick marks that you prefer etc., e.g., *select* the axis, *right click* to select it, *choose* **[Format Axis…]** and then for **Tick Marks**, *select* **[Outside]** for **Major type**. *Check* that your finished graph looks something like the one shown in Fig.1.14.

Once you've formatted the graph into a format that you like the look of, you should save it as a **chart template**. That way you can get the same format again automatically. If you don't already know how to do this, go into Excel **Help** and search on "save chart template". On the companion website, I've also posted a video How to make a scientific graph in Excel using Templates. You can use the preformatted spreadsheet BPM.Ch01_Excel_template.xlsx to *save* the template on your device.

About what you discovered: Change Chart Type and Format Data Series…

In scientific graphs, it's important that you choose the correct **chart type** and **format** for each **data series**. I chose to use Excel's **Scatter with only Markers** for Fig.1.14 because that format indicates that the data are a series of **discrete number pairs** that don't necessarily follow on from the previous value. Most **experimental data points** fit into this same category and should be plotted as markers (or symbols) with each symbol representing the pair of numbers for that data point. These points should *not* normally be connected with lines. However, in some circumstances it's okay to "connect the dots" to make the pattern followed by the data series easier to see (like in your answer to Q.1.3), but this is an exception to the rule that should be used sparingly. Connecting the dots shouldn't be done if the graph also includes a theoretical curve.

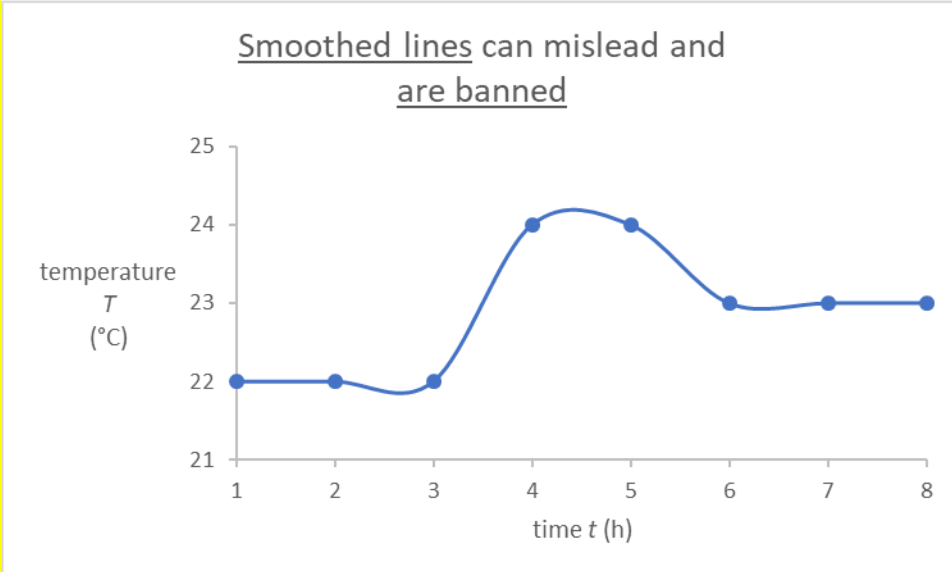


Fig.1.15 Excel **X Y (Scatter)** chart of temperature versus time. The **Data Series** has been **Format**ted as **Scatter with Smooth Lines and Markers**. You should *never use* **Smooth Lines** in your Excel chart answers for questions in this book.

If the data plotted represent a **continuous function** such as a theoretical curve, that curve should be plotted using Excel's **Scatter with Straight Lines**. In your spreadsheet answers, you should *never* use Excel's **Scatter with Smooth Lines**. There are a number of reasons for this. The first is that the smoothing can mislead you as to what's really going on. Consider the graph of temperature as a function of time shown in Fig.1.15, which is based on an example I found on weather.com one day in 2012.

You should note that the **Smooth Line** shows a dip in the temperature during hour 2 even though the data points at hours 2 and 3 are the *same* (22 °C). If we removed the point markers from this graph, things would be even worse! The smooth continuous curve implies that the temperature was measured continuously with an accuracy of better than 0.1 °C (which was simply not the case in this example). Whenever you are formatting a graph, you should think carefully. The format you chose implies something about the data that's plotted. ❑

Q.1.7 *Record* a copy of your graph.
Hint: You should *cut and paste* your Excel graph into your Word document answer. *Click* on the Paste Options icon (Ctrl) (in Word, just after the paste) and *select* the **[Picture (U)]** option. You need to do this to stop the chart changing in Word when you make a change in your spreadsheet. Otherwise, your correct answer can be changed to an *incorrect* one by Microsoft! Also, don't forget to *label* the graph with the question number.

Extend the spreadsheet to 1000 turns. One way to do this on a PC is to *highlight* the row you want to copy and then *press* CTRL+C to select it. *Highlight* the top left target cell (**A103** in my case) and then *hold down* the SHIFT key and, while continuing to hold down SHIFT, *navigate* to the far right-hand corner of your target area using the RIGHT ARROW and the PAGE DOWN keys.

When you get there, *release* the SHIFT key, and then either *press* CTRL+V to copy or simply *press* ENTER. *Extend* the series in your chart to show all 1000 turns. You can *right click* on the graph and then *use* **Select Data > Edit (Series 1)** and *change* the **A102** in the **Series X values:** window to **A1002** and the **B102** in the **Series Y values:** window to **B1002**.

Q.1.8 *Record* a copy of your graph.

Q.1.9 *Select* a blank cell in your spreadsheet and *tap* the DELETE key repeatedly. *Briefly describe* in words what happens to the dots in your graph. Does this behavior remind you of anything?

About what you discovered: mindfulness and discovery

While you're working on all that spreadsheet stuff, it's essential that you keep in mind why you're doing all that work. It's not just to get the spreadsheet working – it's to **discover** for yourself something about how things work. ❑

Q.1.10 **Discussion Question** **(a)** Instead of using Excel and **RANDBETWEEN**, *imagine* that you used a real ten-sided dice to produce a similar graph of the random number r rolled as a function of turn number, then as carefully and as precisely as possible, *describe* how you would expect the dots to be distributed, i.e., how many of each number would you expect to get in your graph (on average) and would you expect to see any consistent patterns?
(b) *Briefly discuss* whether your Excel graph shows this expected behavior.
(c) *Briefly describe* in words how you would quantitatively test your hypothesis about your spreadsheet data.
(d) **Challenge Question** (extra credit) *Perform* the test you outlined in part (c) and *report* your conclusions.

Q.1.11 *Briefly explain* how the two arguments (**Bottom** and **Top**) and the **Formula result** of **RANDBETWEEN** are quantitatively consistent with the graph you produced in Q.1.8. **Hint:** If you need a little help, see the following AWYD.

About what you discovered: using Excel and getting help

Now that we have some idea of what **RANDBETWEEN(1,10)** actually does, let's see what Microsoft has to say… *Select* a blank cell in your spreadsheet and then *left click* on the **Insert Function** icon *fx* to the left of the **Formula bar**. This opens an Insert Function dialog box that you can use to select any of the predefined functions. You can *find* **RANDBETWEEN** by *searching* on the word random, or by *selecting* the **Math & Trig category**. If you've already found **RANDBETWEEN**, then you can *select* the **Most Recently Used category**. When you *click* on **RANDBETWEEN** in the **Select a function:** list, you'll get a brief definition of the function below the list

RANDBETWEEN(bottom,top)
Returns a random number between the numbers you specify.

This description tells you the **name** of the function – **RANDBETWEEN**. It tells you that the function has two **arguments** (**bottom** and **top**). The next line is a very brief definition of the function. You should *read* this definition very carefully to *make sure* that it's actually the function you want. When you *click* on **[OK]**, you'll get the **Function Arguments** window – see Fig.1.16. This also tells you about the function and gives you dialog boxes to *enter* the two **function arguments** **Bottom** and **Top**. When I typed the number **1** in the **Bottom** box and the number **10** in the **Top** box, the **Function Arguments** window looked like Fig.1.16.

Under the dialog boxes there is a very brief technical definition of the function. This definition changes depending on context. You should *read* this definition *very carefully* while you are entering the numbers to make sure you know what every single part of it means *before* you click on **[OK]**. In Fig.1.16, the cursor was in the **Bottom** dialog box and the explanation of what **Bottom** was shown in the middle of the window: **Bottom is the smallest integer RANDBETWEEN will return.** The first part **Returns a random number between the numbers you specify** tells you what number the function returns to the spreadsheet. This is the number that you plotted.
Hint: You should have noted that it was between 1 and 10 in your answer to Q.1.11.

The window also says that the **Formula result = Volatile** (…this doesn't mean that **RANDBETWEEN** is hot-tempered, capricious, or fickle ☺… it just means that you can't predict exactly what value **RANDBETWEEN** will have…). Every time you use a new function, you should read the definition very carefully to figure out what it actually does.

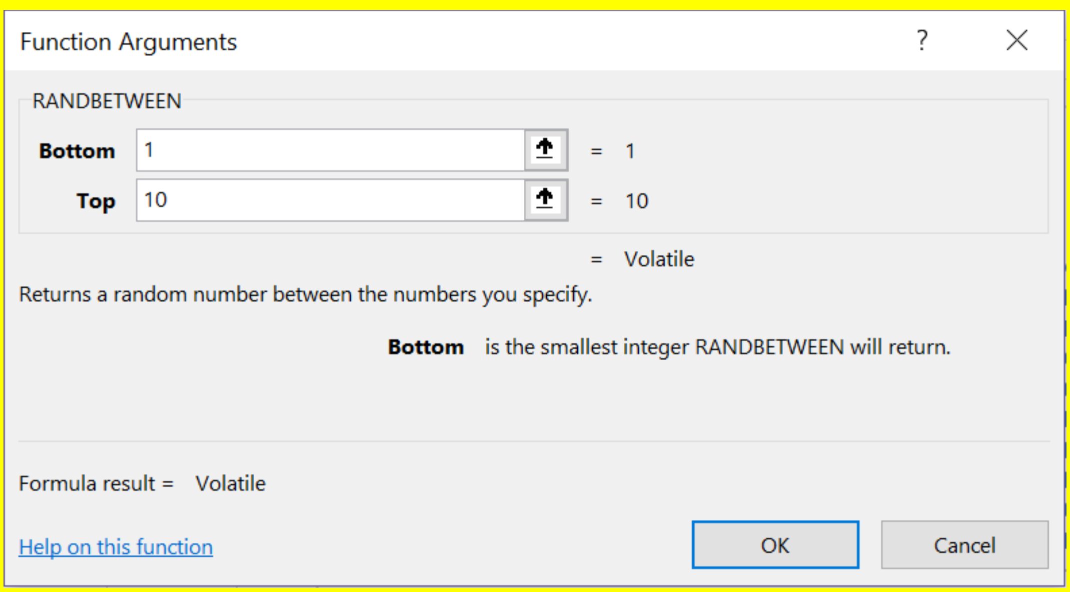


Fig.1.16 Screenshot from Excel. Function Arguments window for **RANDBETWEEN**.

Getting help

By the way, to learn more about this (or any function), you can select **Help on this function** at the bottom of this (the **Function Arguments**) window. The **Help** feature is a very useful source of information about Excel. The **RANDBETWEEN** help page gives examples of how **RANDBETWEEN** can be used. It also tells you **A new random integer number is returned every time the worksheet**

is calculated. There's also a link to **How to copy an example**, which explains how you can copy the example to a blank spreadsheet. In this section of **Help**, there's an explanation of **Formula Auditing**, and the **Formula Auditing Mode**. You can toggle between the regular view of the spreadsheet and the Formula Auditing Mode by pressing CTRL + ` (grave accent), (i.e., hold down the CTRL (Control) key and press the ` (or ~) key). This mode shows the formulas in each cell. It can be very useful during **debugging** a spreadsheet.

Troubleshooting hints

If you had trouble with any of the steps, try going through each one very slowly. The frustrating thing about computers (and Excel in particular) is that the darn thing does exactly what you told it to do, and not necessarily what you meant it to do. The way to check things is simply to go through each step very carefully and check that what you actually told Excel to do – was what you really intended. Left click in the **Formula bar** to check an individual cell or use Formula Auditing Mode to check multiple cells. If you're still having trouble, talk to a friend or a classmate or ask your instructor, read the **Help** screens, or read the related sections again! If you just can't seem to get the spreadsheet to work, then try starting again with a blank spreadsheet. Sometimes this is the most efficient way to fix the problem. It's a little like being told there's a typo or a grammatical error in a paragraph that you've just written. Sometimes trying to find the error is harder than just rewriting it again without copying anything from the original. ❑

1.3 Teaching Excel to play the marble game

Okay, so now that we know how to roll a ten-sided dice in Excel, we're ready to complete our spreadsheet of the marble game. *Save* your spreadsheet for Q.1.8, then *save* a new copy as "BPM.Q.1.12". Turns 2, 3, 4 ... 1000 will just repeat turn 1, so let's *delete* them (just for now) from rows **4** through **1002** of the spreadsheet. Let's also *delete* the chart that's no longer needed. We're now going to complete turn 1 (row **3**). Once it's working correctly, we'll copy row **3** to complete turns 2, 3, 4…

Turn 1 needs to include the game rule about whether N_1 (the number of marbles in box 1) will increase or decrease by one in the first turn. Excel has a function **IF()** that we'll insert into the spreadsheet using the Insert Function icon f_x to the left of the **Formula bar** in a similar manner to what we just did with **RANDBETWEEN()**. *Select* cell **C3** then *click* on the Insert Function icon f_x to the left of the **Formula bar** to open an Insert Function dialog box. You can *find* **IF()** by searching on the word **if**, or by looking in the **Logical category**. If you've already used **IF()**, it will also appear in the **Most Recently Used category**.

When you *click* on **IF** in the **Select a function:** list, you'll get a brief definition of the function (below the list):

> **IF(logical_test, value_if_true, value_if_false)**
> Checks whether a condition is met, and returns one value if TRUE, and another value if FALSE.

This description tells you the name of the function **IF**. Inside the parentheses it lists the three arguments: **logical_test**, which is the condition to be tested; **value_if_true**, which is the value the function returns if the condition tested is **TRUE**; and **value_if_false**, which is the value the function returns if the condition tested is **FALSE**.

Select the **IF** function by *clicking* on the **[OK]** button. The **Function Arguments** window will then open. This window has input boxes for each of the three arguments of the function. You get information about each of them as you put the cursor in each input box. For the cell **C3** that we selected, you should *enter* **B3<=C2** for **Logical_test** (where **<=** is pronounced "less than or equal to"). *Enter* **C2-1** for **Value_if_true** and *enter* **C2+1** for **Value_if_false**. When you *click* on **[OK]**, the function inserted by Excel into cell **C3** should look like **=IF(B3<=C2,C2-1,C2+1)**. *Check* that by *clicking* in cell **C3** and then *clicking* in the **Formula bar**.

	A	B	C	D
1	Turn	*r*	N_1	
2	0		3	
3	1	7	4	
4				

Fig.1.17 Screenshot from Excel.

Highlight cells **A3** through **C3** as shown in Fig.1.17 to *select* all of turn 1, then *copy* all of row **3** to generate a 5-turn game. We now have a "live" 5 turn game… but how can we tell if it's working correctly? One way would be to check that each turn (row in the spreadsheet) obeys the rules of the game. But wait… we already figured out how the game should work in Q.1.2. To use those results to check the spreadsheet, we'll need to temporarily change the *volatile* random numbers (generated by **=RANDBETWEEN(1,10)**) with the *fixed* random numbers listed in Q.1.1. We can do this by simply typing the correct sequence of numbers into the r column.

Q.1.12 *Replace* the random numbers in turns 1 through 5 with the fixed random numbers $r = 6, 4, 4, 7, 6\dots$ (you can simply *type* the numbers in cells **B3** through **B7**). *Record* the spreadsheet in both
(a) Normal Mode and in
(b) Formula Auditing Mode (as discussed in the previous AWYD), and then return the spreadsheet back to **Normal Mode**.
Hint: You can *copy* and *paste* the desired section of the spreadsheet into your Word document and *choose* **[Paste as Picture]**. Your answers should then look something like Fig.1.18. (You may have to *resize* spreadsheet column **C** to see its contents correctly.)

Q.1.12(a)

Turn	*r*	N_1
0		3
1	6	4

Q.1.12(b)

Turn	*r*	N_1
0		3
=A2+1	6	=IF(B3<=C2,C2-1,C2+1)

2	4	3
3	4	4
4	7	5
5	6	6

=A3+1	4	=IF(B4<=C3,C3-1,C3+1)
=A4+1	4	=IF(B5<=C4,C4-1,C4+1)
=A5+1	7	=IF(B6<=C5,C5-1,C5+1)
=A6+1	6	=IF(B7<=C6,C6-1,C6+1)

Fig.1.18 Model answer for questions Q.1.12(a) and Q.1.12(b).

You should *double check* that you typed in the fixed random numbers correctly. Then *check* that Excel is actually playing the game according to the rules. That is, *check* that the values of N_1 exactly match the answer you calculated by hand in Q.1.1. Also, *make sure* that the last cell in the spreadsheet reads *exactly* **=IF(B7<=C6,C6-1,C6+1)**. *Note* that the fixed "random" numbers do not change in Formula Auditing Mode as they are just numbers (i.e., they are not calculated using a formula).

Q.1.13 *Replace* the random numbers in turns 1 through 5 with different fixed random numbers $r = 3, 2, 1, 1, 10\ldots$ then *record* the spreadsheet in both
(a) Normal Mode and in
(b) Formula Auditing Mode.
Hint: *Check* that your answer exactly matches the answer you calculated by hand in Q.1.2.

Once you have confirmed that your spreadsheet is working correctly using the fixed random numbers, *delete* rows **4** through **7** of the spreadsheet. *Change* cell **B3** to **=RANDBETWEEN(1,10)**. *Check* that the value of r in Turn 1 is now be back to its correct form with a *volatile* random number representing the roll of the dice.

Next, *select* all of turn 1 (as shown in Fig.1.17) and *copy* it down to extend the game to 1000 turns. Then, *plot* N_1 versus turn (this means make an **X Y (scatter)** graph with N_1 on the y-axis and turn on the x-axis). You can then use your saved **chart template** to *format* the chart using your saved preferences – it's much easier that way! Once you've *filled in* the title and axis labels etc., *left click* in an empty cell and *press* DELETE about ten times. *Pay attention* to what happens to the N_1 axis. Sometimes its maximum value changes between 10 and 12. This **autoscaling** is normally useful (it makes sure you always see all the data), but because we want to compare one game with the next, we'll want to turn it off (just for now). To do that, *right-click* on the y-axis (maybe twice) and then *select* **[Format Axis…] > [Axis Options]**. Now *type* **10** in the **Bounds** entry box for **Maximum**. In Excel 2016 on a PC this should change the **Auto** setting to a **[Reset]** button. This changes the y-axis **Maximum** from **Auto** to fixed. Now *close* the window and *left click* in an empty cell and *press* DELETE about ten times. If your spreadsheet is working correctly, now would be a good time to *save* a fresh copy.

Okay, so now after some work we've taught Excel how to play the marble game, but what can we do with it? Well, back in 1979 when dinosaurs roamed the earth, and Dan Bricklin (Fig.1.19) was developing VisiCalc (the original spreadsheet program), the idea was to get a computer to automatically recalculate something whenever changes were made to the spreadsheet. Every time we hit DELETE in our spreadsheet, we change 1000 values of N_1 according to the rules of our game. So, what would you like to know about the game? … Our goal is to try to understand how it works … in as much detail as possible! One way to do that is to *look* at the graph you've just made and *try to understand* what's going on by *looking* at it and *seeing* how it changes from game to game when you hit DELETE.

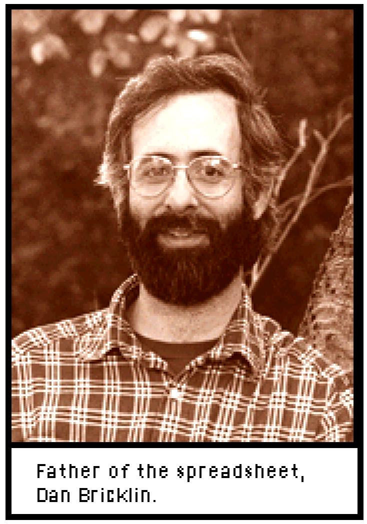


Fig.1.19 Photo of Dan Bricklin from byte.com.

Below are a couple of questions designed to help you think about how the marbles are distributed between the boxes. Just for fun, try to answer them in order without looking ahead. After looking at the raw data for a while (and asking the right questions), you might be surprised by what you can discover just by looking at the graph. Please try to describe what you see initially – no points off for changing your mind later, but please record your initial, carefully considered observation. Before you look carefully at the graph, answer Q.1.14.

> **Q.1.14 DISCUSSION QUESTION** The marble game rule moves the marbles randomly between the two boxes, so N_1 the number in box 1 should also be random. Does this mean that N_1 should be *evenly* distributed (with all values of N_1 being *equally* likely) – just like the numbers r rolled by the digital dice **RANDBETWEEN(1,10)**? *Briefly explain*.

> **Q.1.15** Now *look carefully* at your graph and concentrate on how many times each value of N_1 occurs during the game. *Press* DELETE a few times to get an idea of what usually happens on average. In order to think about this distribution, let's consider a particular number of marbles in box 1. For example, what's the chance of finding $N_1 = 0$? (i.e., no marbles in box 1)? *Look* at your graph and *count* how many times do you see $N_1 = 0$ in your graph, on average. Press DELETE a few times to *estimate* the **observed probability** $P(0)$ of finding $N_1 = 0$ in any given turn.
> **Hint:** A good way to **estimate** the observed probability (i.e., *calculate* a number) is to *count* the number of times you saw $N_1 = 0$ in say 10 (1000 turn) games. If you saw 9

occurrences in 10,000 turns, you'd *divide* 9 by 10,000 to get $P(0) = 0.0009$. The more games you use – the more accurate your answer will be.

Q.1.16 *Briefly describe* in words what the graph tells you about how the probability $P(N_1)$ (the probability of N_1) changes as you go from $N_1 = 0$ to $N_1 = 5$ and then to $N_1 = 10$. Is the probability constant? If not, *briefly describe* how it changes. Look at the graph some more… the answer is there.

Calculating averages is something that baseball fans are often interested in. Excel has a built-in function to do just that. *Type* the heading "**Average**" in cell **D1**. *Click* in cell **D2** and then *click* on the Insert Function icon f_x and *use* the **Help** feature to figure out how to calculate the average. Now, *get* your spreadsheet to *calculate* the average value of N_1 over the whole game. We'll call this value $\langle N_1 \rangle$, where the angle brackets $\langle\ \rangle$ are pronounced "the average value of".

Q.1.17 DISCUSSION QUESTION **(a)** By *pressing* DELETE in a blank cell a few times, *estimate* a good value for $\langle N_1 \rangle$.
Hint: A good answer should only give the answer to the number of decimal places that you think are significant. It should also include an indication of the uncertainty, e.g., something like $\langle N_1 \rangle = 4.9 \pm 0.4$.
(b) *Briefly describe* how you estimated the uncertainty. If you've taken a statistics course… *use* what you learned!
(c) *Record* a typical graph of N_1 versus turn.

About what you discovered: reading graphs

Your answer to Q.1.17(c) should look something like Fig.A1.1 (click the hyperlink to go to the appendix to see it). When you look at this graph of N_1 you might think it looks similarly random to the earlier graph of r **= RANDBETWEEN(1,10)** from Q.1.7 (see Fig.1.14). At first glance the two graphs do appear to be similar, but if you look carefully, you should notice some important differences.

The first major difference is that the graph of N_1 extends over ten times as many turns as the graph of r. After reading the title of the graph, the first thing you should always do is read the axes titles and labels looking for content, scale, and units of the graph. Before you do that, you're only *guessing* what you're looking at. Secondly, while both graphs show integer values for N_1 and r, with a maximum of 10, the N_1 graph (occasionally) shows values of $N_1 = 0$, but r does not. Why is that? Finally, after the analysis you did in Q.1.16, you should know that the probability of N_1 has a maximum at $N_1 = 5$ and decreases to quite small values at $N_1 = 0$ and $N_1 = 10$. In contrast, you should have noticed in Q.1.10 that the probability of r is **uniform** (i.e., the probability of any r in the range from 1 to 10 is $P(r) = 1/10$, a constant value). Finally, you might also notice that the y-axis labels in both graphs are not written vertically (as is customary in Excel and scientific graphs). This is because horizontal labels are much easier to read. This is useful in presentations, and other situations where you want other people to be able to read your graph quickly and easily. ❑

Q.1.18 **Discussion Question** **(a)** Based on your observations, *briefly describe* the shape you would expect for a graph of $P(N_1)$ versus N_1.
1st Hint: If you rotate your graph (or Fig.A1.1) by 90°, you can almost "see" this distribution.
2nd Hint: The shape of this graph is a familiar one, it may even ring a ____!
(b) *Briefly describe* in words how you would quantitatively test your hypothesis.
(c) **Research Question** *Perform* the test you outlined in part (b) and *report* your conclusions.

1.4 Properties of larger marble games

How would you change the spreadsheet in Q.1.17 to play a marble game with $N = 2$ marbles… or $N = 100$ marbles… or… lets go crazy… *one million* marbles? Is it possible? How could you do it? Does much change when you alter the total number of marbles?

Changing the total number of marbles to $N = 100$ in the marble game requires changing every cell in column **B**. To do this, *change* cell **B3** to **=RANDBETWEEN(1,100)**, which corresponds to a hundred-sided dice! Then, *copy* its formula down all 1000 turns. One way to do that is to highlight cell **B3** then *left-double-click* on the bottom right-hand corner blob of the cell when it's highlighted. You should *notice* that the curve that's generated by Excel shoots off the top of the graph because we have a fixed maximum on the N_1 axis. To correct the problem, *set* the scaling of the N_1 axis back to **Auto**. (In Excel 2016 on a PC you can do this by opening the **Format axis** window for the y-axis and in **Axis Options > Bounds** click the **[Reset]** button next **Maximum** to set it back to **Auto**.) You should then *notice* that the graph of N_1 vs. turn now looks quite different! You should also *notice* that the markers are so close together that they can no longer be seen clearly. In this case using markers doesn't make sense because the overlapping markers obscure the details. To correct that problem, *change* the chart to a scatter chart with straight lines and no markers. On a PC you can do this by *right clicking* on the data series in the chart. Select **Format Data Series…** then *click* on the (paint can) icon and *change* **Line** to **Automatic** and *change* **Marker Options** to **None**. Connecting the dots in this manner makes much more sense visually, as more detail can be seen. However, you shouldn't forget that the data are actually just a series of discrete numbers and not a continuous mathematical function.

Q.1.19 *Press* DELETE in a blank cell about ten times to get an idea for how the game works on average and then *record* a representative graph of N_1 *vs.* turn for this 100-marble game that starts with $N_1 = 3$ at turn 0, but first check that the title and axes are still correct. As always, don't forget to *save* your spreadsheet as e.g., "BPM.Q.1.19" once it's working. This is your last reminder about saving spreadsheets as you go.
Hint: If your graph still has **markers** – you didn't *read* the previous paragraph carefully enough!

Q.1.20 DISCUSSION QUESTION **(a)** *Change* the value of N_1 in turn 0 to be 100 marbles and *press* DELETE in a blank cell enough times to get a feeling for what happens on average. *Pay attention* to the value of N_1 after a couple of hundred turns. Now *try* some other starting values in the range $N_1 = 0$ to $N_1 = 100$. After a few hundred turns, the number of marbles in box 1 should approach a consistent value on average. Just by looking at your graphs, *visually estimate* that equilibrium value $\langle N_1 \rangle$.
Hint: In this book, the word **estimate** is used in the same sense as in science and statistics. It means "*find a representative number*."
(b) *Change* the value of N_1 in turn 0 back to 100 marbles. *Look* at the graph and *briefly explain* if the initial decline from $N_1 = 100$ should be included in your estimation of the equilibrium value of $\langle N_1 \rangle$.

About what you discovered: approaching equilibrium

You should have noticed that after a few hundred turns that the number of marbles in box 1 approaches similar values near $N_1 = 50$ marbles. Once at equilibrium, the precise value of N_1 fluctuates, but on average you should be able to convince yourself that N_1 averages out to about $\langle N_1 \rangle = 50$ at **equilibrium**. ❑

Q.1.21 DISCUSSION QUESTION *Briefly answer* the following:
(a) How do you know that the system is at equilibrium?
(b) Is N_1 *always* equal to 50 at equilibrium?
(c) Are the marble game rules any different at equilibrium?
(d) What is happening to an individual marble when the game reaches equilibrium?
(e) What differences are there between the equilibrium of the $N = 100$ marble game and the $N = 10$ marble game?

About what you discovered: equilibrium is dynamic!

After reading standard textbooks, a common misconception is that equilibrium is a static situation where things don't change at all. As we have just discovered, this is not true for the marble game. The rules of the game mean that there are always marbles jumping between the boxes. Molecular equilibrium in physiology is the same – there is always something happening, and equilibrium occurs when the properties of the system stay constant – *on average*. ❑

Q.1.22 CHALLENGE QUESTION Can you think of any other interesting properties of the game that Excel might be able to tell you about? If you can, *briefly describe* how to implement them in Excel.
Hint: For example, how could you summarize the spread in the N_1 values at equilibrium?

Q.1.23 DISCUSSION QUESTION In the original marble game, the marbles are re-labeled at every turn (with marbles 1 though N_1 in box 1). An alternative method would be to permanently assign a number from $1 - 10$ to each marble like the numbers on pool (billiard) balls. The new game rule would then be "Roll the ten-sided dice and move the corresponding marble to the other box," so that if you roll $r = 7$, you move the 7-marble to the other box. In this new game, you can keep track of where each marble is at every turn.

(a) Does this new game rule produce an ***identical*** sequence of jumps with the same marble jumping for a given r? *Briefly explain*.

(b) Does this new game rule produce an ***equivalent*** sequence of N_1 values – in the sense that the new rule has a similar random effect on N_1 – on average? In other words, could you tell the difference between an N_1-versus-turn graph from this new marble game and the original marble game? *Briefly explain*.

(c) *Briefly describe* what additional information about the individual marbles could be extracted using this new marble game.

(d) **CHALLENGE QUESTION** *Briefly describe* how you could implement this game in Excel.

(e) **RESEARCH QUESTION** *Implement* your idea in a spreadsheet and then *present* and *explain* the significance of the additional information.

About what you discovered: marble independence

A fundamental assumption of our marble game model of diffusion is that the marbles don't interact with each other. They simply jump between boxes in a random manner that's independent of where any of the other marbles are located or what they're doing. If all the marbles are in one box, there is no repulsive force pushing them apart or into the other box. Each one moves and jumps to the other box as if the other marbles weren't even there. This idea is explicitly modeled by the alternate marble game of Q.1.23 because any marble (say number 7) is chosen at random in exactly the same way, no matter where the other marbles are. This exemplifies the fundamental assumption of the marble game – that the chance of a particular marble jumping next is always the same. All the marbles are free to jump whenever they "want" – **marble independence**. In this respect, the marble game is a realistic simulation of the *independent* diffusive motion of solutes in dilute solution. That independent motion is quite different from the convective flow of air molecules blown onto the palm of your hand. In convective flow, all the molecules are definitely not independent – they all move together in one direction and, "go with the flow!" ❑

About what you discovered: marble equality

It's easy to see that all marbles are treated equally in the in the alternate "pool ball" version of the marble game (Q.1.23), because if their number comes up on the dice (with a uniform probability of $p = 1/N = 1/10$), then they jump to the other box. It's extremely important to realize that the original marble game (OMG) also treats all marbles equally. The OMG rule is "Roll the ten-sided dice. If r is *less than or equal to* N_1 ($r \le N_1$), then move a marble from box $1 \rightarrow 2$, otherwise move a marble from box $2 \rightarrow 1$." To see how that treats all marbles equally, let's consider a game that currently has $N_1 = 4$ marbles in box 1. The roll of the dice then has a probability of $N_1/N = 4/10$ for satisfying the condition $r \le N_1$. If that condition is true, then there's a uniform probability of $1/N_1 = 1/4$ of selecting any one of the marbles in box 1 to jump to the other box. Hence, the combined probability of selecting a particular marble in box 1 is $p = (N_1/N)(1/N_1) =$

$1/N = 1/10$ just like the pool-ball version of the marble game. Similarly, if $N_1 = 4$, then $N_2 = 6$ and there's a probability of $N_2/N = 6/10$ for the dice roll to *not* satisfy the condition $r \leq N_1$, and a probability of $1/N_2 = 1/6$ of selecting any one of the marbles in box 2 to jump to the other box. Hence, the combined probability of selecting a particular marble in box 2 is $p = (N_2/N)(1/N_2) = 1/N = 1/10$ just like the pool-ball version of the marble game. Hence, the OMG is ***equivalent*** to the "pool ball" marble game because they both use the dice to randomly select the next marble with a uniform probability of $p = 1/N$ – no matter where they are or when they last jumped. ❑

1.5 Physiological applications and implications

Molecular diffusion

In this section, we're going use the marble game to discover how diffusion works. Let's change the total number of marbles in the simulation to $N = 200$. As we discussed in **SECTION 1.4**, you can do that by *changing* **=RANDBETWEEN(1,100)** to **=RANDBETWEEN(1,200)** in all of column **B** of your spreadsheet.

Now, *try* some other starting values in the range $N_1 = 0$ to $N_1 = 200$. You should *notice* that the N_1 graph looks a little smoother and that it takes more turns (or **steps**) for the game to approach the equilibrium value of $\langle N_1 \rangle = 100$ marbles. As shown in Fig.1.20, this allows us to see what's happening during the initial transient period more clearly. After you've finished experimenting, *change* the value of N_1 in turn 0 to be zero marbles.

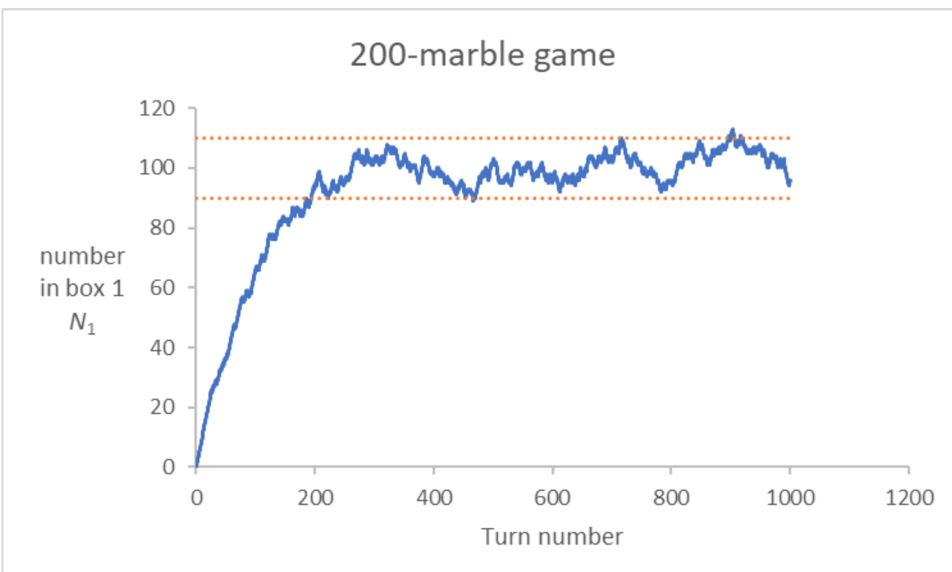


Fig.1.20 Excel **X Y (Scatter)** chart of a 200-marble game starting with all the marbles in box 2. The dotted lines indicate the typical range of N_1 values seen at equilibrium.

Q.1.24 **(a)** With $N_1 = 0$ in turn 0, *press* DELETE multiple times to get an idea of what typically happens in your graph. When you find a graph that has equilibrium values "inside the lines" (similar to Fig.1.20), you should *record* your graph. If you already know how, you should *add* the horizontal lines – if not, we'll learn how to do that in **CHAPTER 2**.

Hint: *Check* that your spreadsheet levels out at about $N_1 = 100$ as shown in Fig.1.20 and don't forget to **[Paste as Picture]** into Word.
(b) *Briefly explain* what the marbles are doing, on average, during turns 1 through 20.
Hint: The only thing that marbles can do is jump between boxes. Hence, a good answer should explain how they're jumping on average.
(c) *Briefly describe* the visual feature of your graph that corresponds to your answer in (b).
(d) *Briefly explain* what the marbles are doing, on average, during turns 100 through 200.
(e) *Briefly explain* what the marbles are doing, on average, during turns 400 through 1000.

About what you discovered: transient *vs.* equilibrium conditions

You should have noticed that when the boxes contain different numbers of marbles, the game rules tend to spread the marbles out evenly between the two boxes in a dynamic process. When box 1 contains more marbles, they tend to move from box 1 → 2, on average. When box 2 contains more marbles, they tend to move from box 2 → 1, on average. A convenient way to quantify the difference between the boxes is to **define** the quantity ΔN (pronounced "delta N") as

$$\Delta N \equiv N_2 - N_1 \tag{1.2}$$

which is the marble **number difference** between the two boxes. The symbol Δ is the uppercase Greek letter "D", which is pronounced "delta". Watch the Greek letters go green! video [Nelson 2013] for a review of the Greek alphabet and review **APPENDIX A.6**. It's standard scientific notation to use Δ as a symbolic prefix to indicate a difference in the thing that it precedes. In our case ΔN is the difference in the N-values (numbers of marbles) for boxes 1 and 2, where N_1 is the number of marbles in box 1, and N_2 is the number of marbles in box 2. The symbol $\equiv$ is an alternate $=$ (equals sign) that can be pronounced "**is defined as**".

One way to think about ΔN is that it summarizes the change in the number of marbles in the current box if you imagine traveling from box 1 to box 2. For example, **Q.** If you go from box 1 → 2 with $N_1 = 40$ and $N_2 = 60$ how does the number of marbles change? **A.** The number of marbles increases by $\Delta N = 20$. **Q.** If you go from box 1 → 2 with $N_1 = 70$ and $N_2 = 30$ how does the number of marbles change? **A.** The number decreases by 40 or $\Delta N = -40$. The minus indicates a decrease going from box 1 → 2. In both cases, equation (1.2) can be used to calculate the marble **number difference** if you go from box 1 → 2. We will be using differences like ΔN often in this book. For example, when the marble game reaches equilibrium, ΔN is approximately zero because $N_1 \approx N_2$ at equilibrium. ❑

We now want to quantify what we've just discovered by considering the net number of marbles that move from box 2 → 1 during the first 100 turns of the simulation. To do that, we'll want to calculate the difference between the values of N_1 at turn 100 and turn 0. *Type* a heading "**Net Transfer**" in cell **E1** of your spreadsheet, then *enter* a formula in cell **E2** to calculate the difference in the N_1 values at turns 100 and 0. This turned out to be **=C102-C2** in my spreadsheet.

Q.1.25 By *pressing* DELETE five (or more) times, *estimate* the average value of this **net transfer** $N(2 \to 1)$, if the value of N_1 in turn 0 is
(a) $N_1 = 0$ marbles,
(b) $N_1 = 50$ marbles,
(c) $N_1 = 100$ marbles,
(d) $N_1 = 150$ marbles, and
(e) $N_1 = 200$ marbles.
Hint: The $N(2 \to 1)$ should be positive values if N_1 increases; and negative values if N_1 decreases – during the first 100 turns of the marble game.

Q.1.26 **Discussion Question** **(a)** *Summarize* your answers to Q.1.25 in a *hand-written* four-column table with headings: $N_1(0)$, the initial (turn 0) number in box 1; $N_2(0)$, the initial (turn 0) number in box 2; $\Delta N(0)$, the initial difference in the number of marbles {where $\Delta N(0) = N_2(0) - N_1(0)$}; and $N(2 \to 1)$, the average net transfer during the first 100 turns. *Record* a picture of your hand-written answer.
1st Hint: The values for $N_2(0)$ are $200, 150, 100, 50, 0$ for parts (a), (b), (c), (d), (e) respectively so that the values for $\Delta N(0)$ are $200, 100, 0, -100, -200$ for parts (a), (b), (c), (d), (e) respectively.
2nd Hint: The $N(2 \to 1)$ column numbers should correspond to your answers in Q.1.25(a)-(e).
(b) *Briefly explain* how the values for the initial difference $\Delta N(0)$ relate to the initial number in box 1 $N_1(0)$. Your answer can be a mathematical equation.
(c) *Briefly describe* how the values for the average net transfer $N(2 \to 1)$ relate to the initial difference $\Delta N(0)$.

About what you discovered: jump from high to low

In your answer to Q.1.25(a), you estimated an average net transfer of $N(2 \to 1) \approx 60$, i.e., a number somewhere in the range of 50 to 70 marbles. $N(2 \to 1) = 60$ means that 60 marbles were transferred on average from box $2 \to 1$ during the first 100 turns of the marble game when the marble game started out with $N_2 = 200$ marbles in box 2 and $N_1 = 0$ marbles in box 1. This can be summarized by saying that the marbles jump from **high to low** on average. Similarly, in Q.1.25(d) you estimated an average net transfer $N(2 \to 1) \approx -30$, i.e., a number somewhere in the range of -40 to -20 marbles. $N(2 \to 1) = -30$ (note the minus sign) means that 30 marbles were transferred on average from box $1 \to 2$ during the first 100 turns of the marble game when the marble game started out with $N_1 = 150$ marbles in box 1 and $N_2 = 50$ marbles in box 2. This can also be summarized by saying that the marbles jump from **high to low** on average. ❑

Q.1.27 **Discussion Question** **(a)** In light of what we discussed in the previous AWYD, *briefly discuss* whether all the data in the table you made for Q.1.26(a) can be summarized by "the marbles jump from **high to low** on average." Don't forget to talk about the data for Q.1.25(c).

(b) The quantity $\Delta N(0)$ summarizes quantitatively which box is high and which one is low, e.g., $\Delta N(0) = 100$ means that box 2 starts out high and box 1 starts out low, whereas $\Delta N(0) = -100$ means that box 1 starts out high and box 2 starts out low. $\Delta N(0) = 0$ means that the boxes start out even. We're now going to investigate the relationship between $\Delta N(0)$ and $N(2 \to 1)$ by making a graph – that's what scientists do! In order to do that, *enter* your table from Q.1.26 into a *blank* spreadsheet, i.e., a new Excel file, and use it to *plot* a **Scatter with only Markers** chart that summarizes how the net transfer $N(2 \to 1)$ during the first 100 turns depends on the initial difference in the number of marbles $\Delta N(0)$. *Record your graph*.

(c) By looking at your graph from part (a), *briefly summarize* in words what you can conclude about how the difference in the initial number of marbles $\Delta N(0)$ influences the net transfer $N(2 \to 1)$ of marbles (during the first 100 turns).

About what you discovered: Fick's law of diffusion

What we've just discovered is one of the most important general principles in biophysics and physiology! If there are two similar regions (think boxes) that molecules can jump between via *unbiased* random Brownian motion, then the molecules will tend to move from regions with high concentration to regions with low concentration, on average. This average motion from **high to low** that occurs as a result of unbiased random Brownian motion is called **diffusion**. According to a 2011 survey conducted by the Association of American Medical Colleges (AAMC), "Transport Processes" is the second most important topic in medical education after "Nucleic Acids." As we've just discovered, the marble game realistically simulates the most important transport process – **diffusion**. In the following section, and in later chapters, we'll use the marble game to discover how many other physiological processes operate.

Your answer to Q.1.27(a) should look something like Fig.A1.2 and have $\Delta N(0)$ on the x-axis and $N(2 \to 1)$ is on the y-axis. This graph shows that the **net transfer** $N(2 \to 1)$ is directly **proportional to** the **initial difference** $\Delta N(0) = N_2(0) - N_1(0)$. That observation is known as **Fick's law of diffusion**. The process we followed is a basic **scientific method** that scientists use to learn about the physical world: **(a)** Perform an experiment to measure quantities that appear to be related, e.g., net transfer $N(2 \to 1)$ and initial difference $\Delta N(0)$. **(b)** Plot those quantities to discover if there is a relationship between them.

In my graph (Fig.A1.2), I added a linear trendline to visually validate the relationship that they are proportional. We'll learn how to do that kind of thing in **CHAPTER 4**.
Note: The trendline isn't required for your answer to Q.1.27(a).

Our marble game is a **simulation** of what happens in the real world. Simulations are an extremely useful abstraction in science because *we* get to decide *exactly* what the rules are. Anything that happens in the simulation must therefore be a consequence of our rules. As a result, we can conclude that Fick's law is a *natural consequence* of the random Brownian motion that is simulated by the jumps in the marble game.

By performing numerous experiments in the real world, scientists have discovered that many molecules in physiology (and in other branches of science) move from place to place in a manner that can also be described, predicted, and understood using Fick's law of diffusion. ❑

Q.1.28 **DISCUSSION QUESTION** **(a)** *Briefly describe* what the marbles do in the marble game that causes Fick's law of diffusion in Fig.A1.2 and your answer to Q.1.27(b)?
(b) Based on the physical basis for the jumps in the marble game (Fig.1.4), *briefly summarize* the actual physical process that causes Fick's law of diffusion.

Q.1.29 **RESEARCH QUESTION** By doing a literature search, find two examples of quantitative scientific evidence for Fick's law of diffusion. *Briefly describe* the experimental methods, results, and conclusions of what you found – don't forget to cite references.

About what you discovered: Brownian motion causes diffusion

The marble game is an authentic simulation of molecular diffusion in biophysics and physiology. It accurately demonstrates that random Brownian motion (e.g. as shown in Fig.1.4 and the Brownian motion demo spreadsheet and the Web VPython program MarbleGame) as simulated by jumps in the marble game is *the cause* of Fick's law of diffusion. Marbles jump at the same rate between boxes *independently* of how the marbles are arranged. The spreading out is caused by pure chance (randomness). No physical **force** (a push or pull) is required to spread the marbles out evenly between the boxes. Unbiased random jumps, caused by random thermal motion, are all that are needed to produce diffusion from high to low, as shown in your answer to Q.1.27(b) and Fig.A1.2. ❑

Q.1.30 **DISCUSSION QUESTION** In light of what you've discovered from the marble game, *briefly summarize* how diffusion is different from convection.
Hint: Review the "diffusion *vs.* convection" AWYD and the "marble independence" AWYD after Q.1.23.

Q.1.31 **DISCUSSION QUESTION** There are many misconceptions surrounding diffusion and what causes it to occur. How would you use what you've discovered from the marble game to *explain what's really going on* to a student who had the following misconceptions?
(a) "*Molecules are pushed from high to low concentration*".
Hint: Is there really a physical force pushing the molecules from high to low?
(b) "*Molecules go where they are needed*".
Hint: How would molecules know where to go? They're not conscious beings!
(c) "*There is less force on the molecules when they are less crowded*".
Hint: See the hint for (a).
(d) "*Going from high to low gives the molecules more space*".
Hint: Is a molecule really like someone in a bad relationship?
(e) "*When there is no concentration gradient, the molecules don't move*".

Hint: Does Brownian motion (jumps between boxes) stop when the system reaches equilibrium?
(f) *Briefly describe* any other misconceptions about diffusion that you've corrected or heard from others?

Physiological implications

Now that we've discovered that the marble game produces diffusion according to Fick's law, let's see how it can be applied to physiological systems. Fig.1.21 shows how the marble game can model an important physiological situation – O_2 delivery from blood to heart muscle tissue. In this system, O_2 molecules jump between the boxes by random Brownian motion similar to that shown in Fig.1.4, the Excel Demo spreadsheet, and the Web VPython program MarbleGame. Our marble game can be used to **model** this system because the two boxes are very similar salty solutions (at least from an O_2 molecule's perspective). Box p contains blood plasma in a cardiac capillary, and box c contains cytosol in an adjacent heart muscle cell (cardiomyocyte). The blood plasma and cytosol boxes are very tiny and contain no large structures like red blood cells or mitochondria. This means that jumps in both directions are equally likely. In this situation, the marble game is a very good simulation of the real transport process. As a result, we can use what we've learned from the marble game to predict how the numbers of O_2 molecules in each box (N_p and N_c) are related.

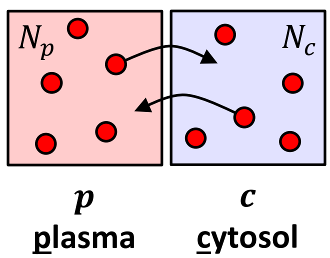


Fig.1.21 Marble game representation of O_2 molecules moving between two tiny boxes. Box p contains blood plasma in a capillary and box c contains cytosol in an adjacent cardiomyocyte. N_p is the number of O_2 molecules in box p, and N_c is the number of O_2 molecules in box c.

Q.1.32 **(a)** For Fig.1.21 to represent normal physiological conditions for a healthy person (with a net transfer of O_2 molecules from box $p \rightarrow c$), can the marble game be at equilibrium? *Briefly explain*.
(b) If not, *briefly explain* what must be the mathematical relationship between N_p and N_c? (e.g., is $N_p < N_c$?)

About what you discovered: equilibrium means death!

For an aerobic organism, O_2 equilibrium would result in *no* net transfer of O_2 to tissues resulting in death! Even when you hold your breath, this does not happen. Blood plasma always has a higher O_2 concentration than tissue. ❑

Diffusion enhancement

In mammals, the higher O_2 concentration in box p (in Fig.1.21) is maintained by a store of O_2 (outside of the plasma) that's bound to hemoglobin (Hb) molecules trapped inside red blood cells. Hb is the molecule responsible for the color of these blood cells and whole blood. When O_2 molecules are delivered to tissue, Hb releases fresh O_2 molecules, keeping the concentration of O_2 molecules higher in the plasma and maintaining $N_p > N_c$ throughout the circulatory system.

Q.1.33 If the marbles in Fig.1.21 represent CO_2 molecules (instead of O_2 molecules), what must be the mathematical relationship between N_p and N_c in a living organism?
Hint: CO_2 is a metabolic waste product that must be removed from the cytosol.

Q.1.34 **DISCUSSION QUESTION** **(a)** What must be the mathematical relationship between N_p and N_c, if the marbles in Fig.1.21 represent glucose molecules that are consumed in the cell?
(b) Once glucose molecules enter the cytosol, they are transformed (during glycolysis) into molecules that cannot participate in the glucose marble game. *Briefly explain* how this enhances diffusion of glucose to tissue.
(c) Some people might say that "glycolysis is a process that pulls glucose into the cell". *Briefly explain* why the word "pulls" might lead to misconceptions about what is really going on.

Q.1.35 **DISCUSSION QUESTION** Fetal Hb is different from maternal Hb. In order for O_2 to diffuse from mother to fetus (i.e., from maternal plasma to fetal plasma – across the placenta epithelia), there must be an O_2 concentration difference between maternal plasma and fetal plasma.
(a) *Sketch* (*draw by hand*) a marble a marble game similar to Fig.1.21 showing O_2 diffusion between maternal plasma and fetal plasma boxes.
(b) Using your figure, *briefly explain* whether fetal Hb should bind O_2 at higher or lower concentrations than maternal Hb to make O_2 diffuse from mother to fetus.

About what you discovered: glycolysis and fetal Hb

The reason why glycolysis in the box c enhances diffusion into the cytosol is because it reduces N_c to low values, making the difference between N_p and N_c larger. That makes the net transfer from high to low occur more quickly than it would otherwise. The reason why Fetal Hb should bind O_2 at a lower concentration (in fetal plasma) is so that the concentration of O_2 in fetal plasma is lower than in maternal plasma. That way, O_2 will diffuse from high to low – from mother to fetus (while storing more O_2 in the fetal Hb). ❑

During muscle contraction, blood flow to muscle may not be high enough to meet the demand for oxygen. Under these transient conditions, mitochondria continue to consume O_2 – see Fig.1.22. This decreases the O_2 concentration in box m, and O_2 molecules diffuse from box $c \rightarrow m$ lowering the concentration in box c. As a result, the cells in muscle tissues contain another heme-containing protein – myoglobin (Mb) that's responsible for the color of red meat. Mb reversibly binds O_2 at concentrations that are much lower than those in blood. When the O_2 concentration in box c reaches this low value, Mb releases its stored oxygen into the cytosol, thereby maintaining the O_2 concentration in box c at this reduced (but non-zero) level. This ensures that there is a higher O_2 concentration in the cytosol near mitochondria, so that O_2 molecules continue to enter mitochondria (on average, by diffusion) as shown in Fig.1.22.

The operation and functioning of myoglobin and hemoglobin are still an active topic of scientific research and the physiological functioning of myoglobin is still not completely understood. Myoglobin molecules also undergo Brownian motion and those containing O_2 provide a parallel mechanism for delivery of O_2 to mitochondria by playing their own (independent) marble game inside the cytosol. We'll return to this fascinating topic in **CHAPTER 6**.

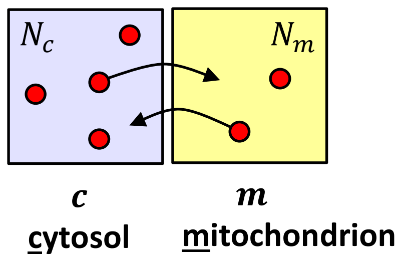


Fig.1.22 Marble game representation of O_2 molecules jumping between two tiny boxes. Box m represents a small piece of a mitochondrion and box c contains cytosol just outside the mitochondrion. (This box c is different from the one shown in Fig.1.21.) N_c is the number of O_2 molecules in box c and N_m is the number of O_2 molecules in box m. O_2 molecules in box m are either consumed or return to the cytosol.

About what you discovered: the oxygen cascade

The steady decrease in oxygen concentration that we've inferred from the marble game is part of the **oxygen cascade**. The oxygen cascade summarizes how the concentration of O_2 molecules must decrease as they travel on average from the lungs through the circulatory system into tissue and finally to the mitochondria where, during oxidative phosphorylation, they combine with hydrogen atoms to form water molecules and release energy. ❑

Conclusion – about what you discovered

Congratulations! If you made it here, then you should be in good shape for the rest of the book! One of the things you might have noticed about our approach is that we rely heavily on reading graphs to figure out what's going on. Most scientific data are presented in the form of graphs. As the saying goes, "a picture is worth a thousand words" … in this chapter we generated graphs that literally contained a thousand data points. Learning how to interpret graphs carefully is an essential scientific skill. It's important for us to be mindful that we're trying to **discover** new insights into how models work and what they're telling you. That way we can **discover science** by **just doing it**.

In this chapter, we taught Excel how to play the marble game. This is a **computer simulation** of the marble game. Simulations like this are an important part of cutting-edge scientific research.

Our simulation (**sim**) behaves just like the real game. *Any* question you might want to ask about the game can be answered using the sim. Simulations that use random numbers are called **Monte Carlo** methods. The name "Monte Carlo" comes from the original casino in the principality of Monaco (not the newer one in Las Vegas), where chance also determines the outcome. A sim can tell us much more than how the game works on average. The randomness inherent in the actual game is realistically reproduced in the sim. Molecular-level systems have similar randomness that can have significant consequences. For example, in neurophysiology the randomness associated with a small number of ion channel proteins can be amplified to form nerve impulses… In the following chapters, we'll modify our marble game to quantitatively model more realistic systems, like those discussed in **SECTION 1.5** of this chapter.

In this chapter, we've discovered an important aspect of the marble game – it reaches a **dynamic equilibrium** that's brought about by the jumping of the marbles from box to box. At equilibrium, the **average number** of marbles in box 1 $\langle N_1 \rangle$ is approximately constant, even though the actual value of N_1 fluctuates. The same equilibrium $\langle N_1 \rangle$ is reached independent of how many marbles started out in box 1 – so long as the marble game has run for long enough that the starting value no longer influences the current value of N_1. Equilibrium is not a static situation – the marbles continue to jump between the boxes.

The marble game simulates molecules jumping between two adjacent boxes. In physiological systems, these **jumps** are produced by the **Brownian motion** of the molecules that is in turn caused by the random thermal motion of the surrounding water molecules. Even though there is no bias in the individual jumps, we've discovered that the marble game tends to move molecules from regions with high concentration to regions with low concentration in a process that's called **diffusion**. We've also discovered that the rate at which marbles are transferred from high to low (on average) is proportional to the concentration difference. This observation is known as **Fick's law of diffusion**.

Using the marble game, we were able to *predict* that physiological systems are not at equilibrium. Concentration differences must be maintained to ensure the net delivery of nutrients (e.g., O_2 and glucose) to cells by diffusion. Hemoglobin and myoglobin store reserve O_2 molecules that are released to maintain the required concentration differences between adjacent regions in the oxygen cascade. This ensures that there's a steady supply of oxygen all the way to mitochondria where the oxygen is consumed. Glycolysis in the cytosol removes glucose molecules from the glucose marble game increasing the concentration difference with adjacent plasma. This *enhances* glucose transport into cells. These transport phenomena are so fundamental to physiology that "Transport Processes" was ranked second only to "Nucleic Acids" in a 2011 survey of medical schools. In later chapters, we'll continue to use the marble game to discover more about these **transport processes**.

Summary: Introduction – the marble game

Key computational concepts

Marble game

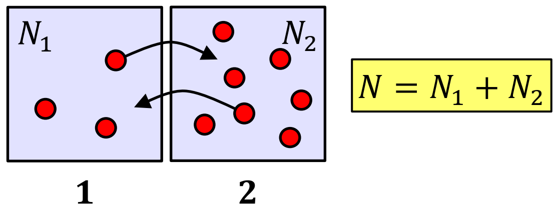


Fig.1.5 Marble game representation of the marble game (repeated from main text).

- The marble game is a **Monte Carlo simulation** (**sim**) that uses random numbers to **simulate** the random Brownian motion of molecules in physiology (e.g., O_2).
- N_1 is the number of marbles in box 1
- N_2 is the number of marbles in box 2
- N is the total number of marbles in the marble game, it's a constant related to N_1 and N_2 by the **bookkeeping equation**

$$N = N_1 + N_2 \tag{1.1}$$

- ΔN is the marble **number difference** between the two boxes, which is *defined* by

$$\Delta N \equiv N_2 - N_1 \tag{1.2}$$

- r is the number rolled by the ten-sided dice, or Excel's **RANDBETWEEN(1,10)** function

Marble game rules

a) Place 3 marbles in box 1 and 7 marbles in box 2, so that $N_1 = 3$.
b) Roll the ten-sided dice. If the number r you rolled is *less than or equal to* N_1, then move a marble from box $1 \rightarrow 2$, otherwise move a marble from box $2 \rightarrow 1$.
c) Repeat step b) using the new value of N_1.

As a result, each marble has an equal chance of jumping next no matter what the value of N_1 is – **marble independence**.

Key physiology concepts

- The marble game reaches **equilibrium** when the value of N_1 reaches a consistent value $\langle N_1 \rangle$, on average.
- **Equilibrium** is a **dynamic process**. At equilibrium there are fluctuations because the molecules continue to jump randomly between the boxes due to random **Brownian motion**.

- The unbiased **jumps** produced by Brownian motion tend to spread the molecules out evenly between the boxes so that they **diffuse** from high to low concentration, on average.
- Using our marble game, we discovered that the rate of diffusion (into box 1) is **proportional to** the marble number difference ΔN. This observation is known as **Fick's law of diffusion**.
- **Fick's (1st) law of diffusion** – molecules diffuse from high to low concentration (on average) at a rate that's proportional to the concentration difference.
- **Hemoglobin** and **myoglobin** store reserve O_2 molecules that are released to maintain a concentration difference between key points in the **oxygen cascade**.
- **Glycolysis** lowers the glucose concentration in cytosol, maintaining a large glucose difference with adjacent blood plasma.

Key probability concepts

- **Observed probability** $P(N_1)$ is the probability of state N_1. I.e., number of occurrences observed in the sim divided by the total number of turns.
- **Observed average** $\langle N_1 \rangle$ is the average of all the N_1s observed during a selected portion of the sim (e.g., after equilibrium is reached).
- In contrast, **expected values** are those predicted by theory – see the following chapters.

Key Excel concepts

- **Formula bar** – found next to the insert function icon f_x. Left click in a spreadsheet cell and then in the **Formula bar** to check the cell formula.
- Excel **functions**

  =A2+1

  uses relative cell addressing to add 1 to the cell above (**A2** in this case)

  =RANDBETWEEN(1,100)

  generates a **uniformly random** integer in the range **1** to **100** (inclusive)

  =IF(B3<=C2,C2-1,C2+1)

  test if **B3** is less than or equal to **C2** and return **C2-1** if it is, and **C2+1** if it isn't

  =AVERAGE(C2:C1002)

  return the **average** of the values in the range **C2:C1002**
- **Relative cell referencing** - Excel default for formulas like **=A2+1**. The relative relationship between the cells is maintained during a copy. Relative addressing allows us to copy a whole row of the spreadsheet thousands of times by clicking and dragging or left-double-clicking in the bottom right-hand corner.
- **Debugging** – Unfortunately, things don't always work correctly the first time! When this happens, you'll need to identify where the problem is. Use the **Formula bar** to check an individual cell and **Formula Auditing Mode** (press CTRL + `) for the whole spreadsheet. You should start at the beginning and make sure that each cell is working properly.
- **File > Save As** – Save your progress to a new spreadsheet filename for each new question. Help out your future self so that they can find your current work later!

Scientific graphs in Excel

- **XY(Scatter) Charts** – Use this format, and *not* a **Line** chart.
- **Discrete data** (e.g., experimental results) should be plotted as **Scatter with only Markers**.
- **Continuous data** (e.g., theory curve) should be plotted as **Scatter with Straight Lines**.
- **Chart axis options** – Select **[Format Axis…]** to change **Axis Options** from **Auto** to fixed. **[Reset]** back to **Auto** to make sure you see all data.
- **Chart templates** – Once you have a chart in a format that you think you'll want to use again, save it as a template so that you can make another like it more quickly.
- You can use preformatted spreadsheet BPM.Ch01_Excel_template.xlsx to *save* a **markers_only** template on your device.
- Press DELETE, or F9 on a P.C. to update, on a Mac use ⌘ + **=**.
- **[Paste as Picture]** into your Word document answer.
- **Smooth lines** – are banned!

Reading scientific graphs

Plotting and interpreting graphs is an important part of being a scientist.

1. Read the title/caption/legend.
2. Read the axis labels and units.
3. Carefully inspect the data.
4. Does the graph make sense?
5. What is it telling you?

Scientific method

Fig.A1.2 summarizes our rediscovery of **Fick's law of diffusion** using graphical methods. This is an example of using a **scientific method** used to discover the **empirical relationship** (Fick's law) that the **net transfer** $N(2 \rightarrow 1)$ is directly **proportional to** the **initial marble number difference** $\Delta N(0) = N_2(0) - N_1(0)$. The relationship is **empirical** because it is based solely on our observations of the marble game.

Mindfulness and discovery

In this book you'll spend a great deal of time working through questions that ask you to do things in spreadsheets. It's always important to be **mindful** of the reason why you're doing all that spreadsheet stuff. It's not just to get the "right" answer for your homework – it's for you to gain insights into how the models work and what they're telling you. That way you can **discover science** by **just doing it**.

Appendix - Answer graphs

Figure A1.1 – Reading graphs

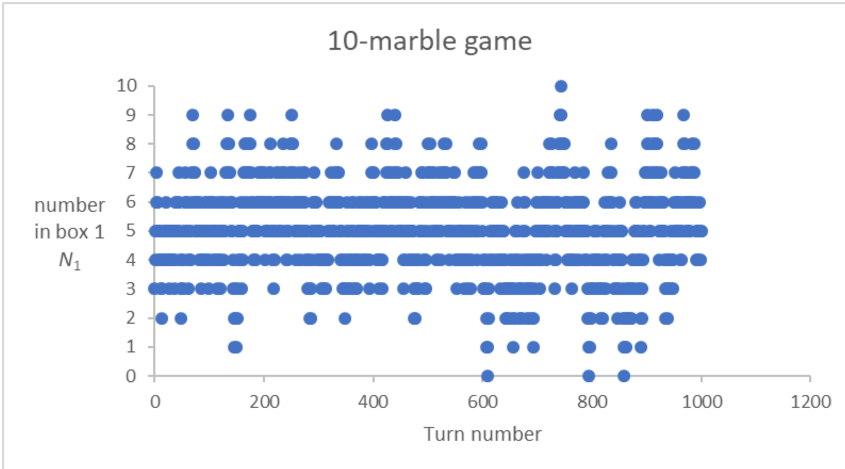


Fig.A1.1 Excel **X Y (Scatter)** chart of N_1 the number of marbles in box 1 for a 1000-turn $N = 10$ marble game.

Return to main text

Figure A1.2 Diffuse from high to low

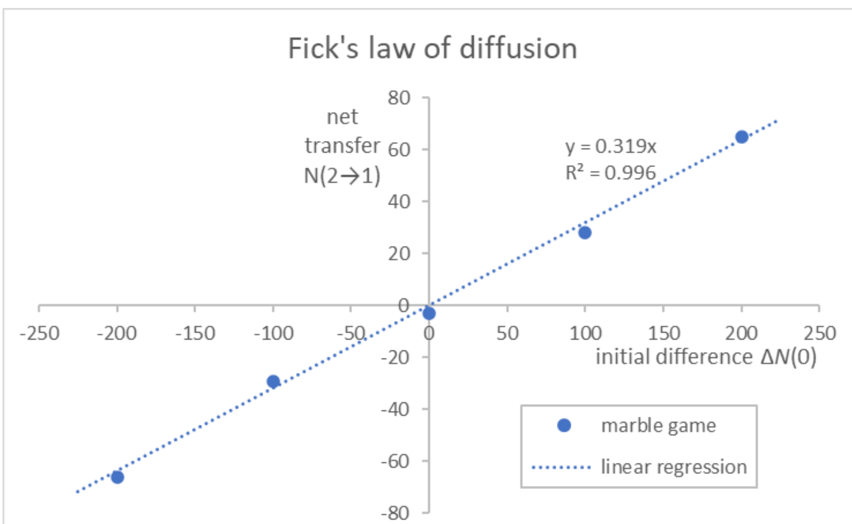


Fig.A1.2 Excel **X Y (Scatter)** chart answer to Q.1.27 showing that the net transfer of marbles from box 2 → 1 during the first 100 turns of a 200-marble game is proportional to the difference in the initial number of marbles $\Delta N(0) = N_2 - N_1$. This is **Fick's law of diffusion**. **Note:** You were not expected to add the dotted trendline – we'll be learning how to do that in **CHAPTER 4**.

Return to main text

References

Nelson, P. H. (2013) *Greek letters go green!* https://circle4.com/biophysics/videos/
Nelson, P. H. (2020) *MarbleGame* Web VPython simulation
https://www.glowscript.org/#/user/LyonSays/folder/BPM/program/MarbleGame

Incoming transmission…

…we are Microsoft…

…you will be **assimilated**…

…resistance is futile! ☺

This material is based upon work supported by the National Science Foundation under Grant Nos. 0836833, 1817282, and 2306506. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.