# Randomness: Its Importance in Statistical Inference and How to Teach Students About It

**Dr Nathan (Norm) Hoffman**

- I want to examine the following problem:
- Many otherwise able tertiary students struggle to master basic concepts in statistical inference.
- In my view the reason for this seems to be that the students have no feeling for the nature of randomness in experimental science.
- If this is so, what can be done to improve the situation?
- My answer is to give students practical experiences with Monte Carlo simulation.
- As a preliminary to further comments consider the following two experiments.

- i) Stand on a smooth, hard surface, say a concrete floor and hold a golf ball at nose level.
- Now release the golf ball and observe the height to which it rebounds.
- Let’s suppose it rebounds to chest height.
- If you repeat the experiment, with the same initial conditions, the ball will rebound to the same height.
- This experiment involves a deterministic process.

- ii) Hold a die (singular of ‘dice’) at about 50cm above a soft surface.
- Now release the die and observe the uppermost face that is showing.
- It will be a 1,2,3,4,5 or 6.
- Suppose it is a 4.
- If you repeat the experiment the result might be a 4, but it is just as likely to be a 1,2,3,5 or 6.
- This experiment involves a random (or stochastic) process.
- The rest of this paper is devoted to random processes.

- In relation to arithmetic, it is well known that involving young children in activities with concrete materials can form a basis for sound understandings about numbers, and the operations with them ( +, -, ×, ÷).
- The same appears to be true with children's understandings about randomness.
- But what does this mean in practice?
- It is common for children to play games that involve chance.
- Monopoly, Snakes and Ladders, and Ludo are examples of such games (played with dice as the generators of random outcomes).
- In spite of this, most students do not, as a consequence, develop sound understandings about randomness.
- How can this be so?

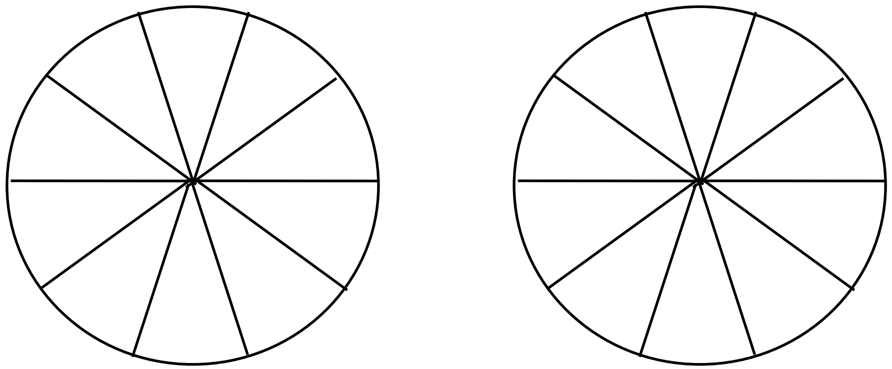
- The ingredient that is missing from the experiences of children playing Monopoly is **the systematic recording and analysis of the outcomes of the chance process.**
- When this is added into the children’s experiences then they are much more likely to develop sound understandings about randomness.
- The example which follows is a Monte Carlo simulation that I have used with 11- and 12- year old's over the past twenty-five years.
- It is an adaptation of a problem given by Arthur Engel (Engel, Arthur. “The Teaching of Probability in the Intermediate Grades.” in “The Teaching of Probability and Statistics”, edited by Lennart Rade, pp87-150, 1970)
- The example, **as presented to the school students**, is as follows:

- **FAMOUS OLYMPIANS**
- Famous Olympian cards are produced by Kolleggs, the makers of the breakfast cereal, Brekky Bites.
- There are 100 different olympians in the whole series.
- Kolleggs put *5 cards in each large packet* of Brekky Bites.
- You can also *buy the cards in packs of 50.*
- One million cards of each olympian were printed.
- The whole supply of cards was then thoroughly mixed before the cards were put in the packs and into the packets of Brekky Bites.
- Imagine that you buy a 50-card pack of the cards:
- 1 Would you expect to get one of each of the 100 different cards?
- 2 Would you expect to miss out on some olympians?

- 3 Would you expect to get exactly one card for some olympians?
- 4 Would you expect to get two cards for some olympians?
- 5 Would you expect to get three cards for some olympians?
- Now imagine that you open your pack and examine its contents:
- a) How many of the 100 olympians would you expect to be 0-carders, ie olympians for whom you have 0 cards?
- b) How many of the 100 olympians would you expect to be 1-carders, ie olympians for whom you have 1 card?
- c) How many of the 100 olympians would you expect to be 2-carders, ie olympians for whom you have 2 cards?
- d) How many of the 100 olympians would you expect to be 3-carders?

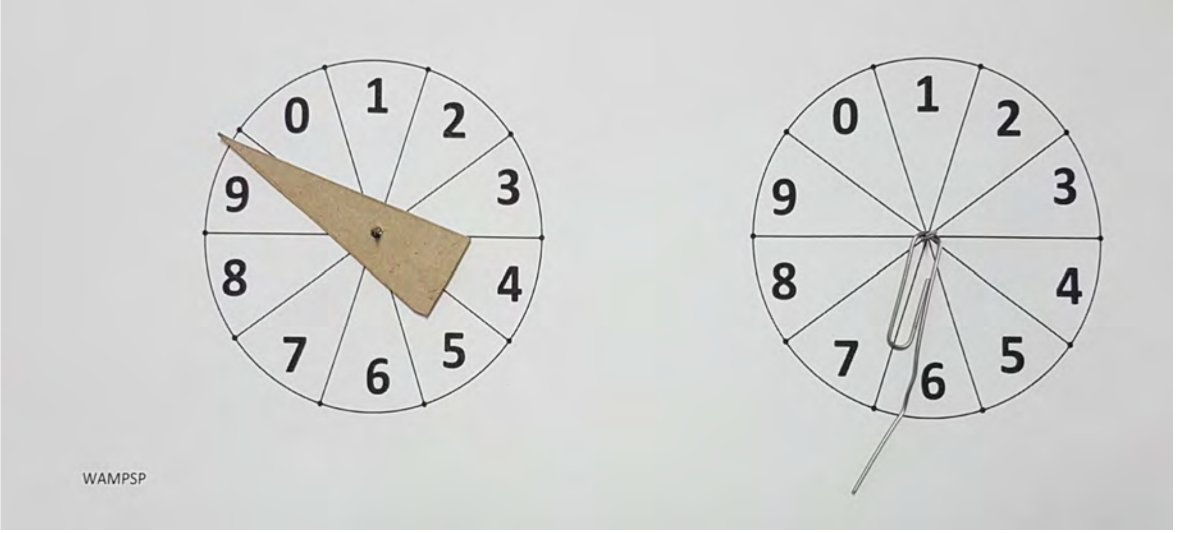
- e) How many of the 100 olympians would you expect to be 4-carders or more?
- We now make a table showing each student’s results.
- We then calculate the class average for a), b), c), d), and e).
- What would be your "best guesses" for the class averages?
- How might we systematically explore this problem?

- **The Famous Olympians Problem – Getting Started**
- 1 Use the sheet that has two "decimal wheels" on it, i.e. two circles, each of which is divided into ten congruent sectors:



- 2 The sectors on each circle have been marked with the digits 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9.
- The circle on the left is for the tens digit, the other for the ones digit.
- 3 Push a thumb tack through the centre of each circle, then turn the sheet over and put the two thumb tacks through the holes, so that the points are sticking up.

- 4 Put the sheet on a **flat surface** and put a spinner on each thumb tack.
- I've shown two different types of spinners.




- 5 Give each spinner a "push" (not "flick") to produce an outcome on each spinner.
- If the digit on the left wheel is 9 and the digit on the right wheel is 6, this counts as being a card for Olympian 96.

- 6 There are two problems with this.
- First, what happens when both spinners show a zero?
- Second, how can you get a card for Olympian 100?
- Problem 1 solves problem 2 – **A double zero is to be counted as a card for Olympian 100.**
- 7 Do **50 double spins and record the outcomes in the table 'Simulation 1'.**
- An example of a completed Outcomes Sheet is shown below.

- **Simulation 1**

| Card | Footballer | Card | Footballer | Card | Footballer | Card | Footballer | Card | Footballer |
|---|---|---|---|---|---|---|---|---|---|
| 1 | *13* | 11 | *22* | 21 | *34* | 31 | *26* | 41 | *99* |
| 2 | *91* | 12 | *39* | 22 | *91* | 32 | *14* | 42 | *78* |
| 3 | *9* | 13 | *54* | 23 | *45* | 33 | *95* | 43 | *66* |
| 4 | *20* | 14 | *77* | 24 | *29* | 34 | *83* | 44 | *42* |
| 5 | *34* | 15 | *97* | 25 | *83* | 35 | *14* | 45 | *41* |
| 6 | *16* | 16 | *90* | 26 | *76* | 36 | *59* | 46 | *27* |
| 7 | *91* | 17 | *95* | 27 | *27* | 37 | *57* | 47 | *15* |
| 8 | *21* | 18 | *99* | 28 | *1* | 38 | *29* | 48 | *37* |
| 9 | *13* | 19 | *22* | 29 | *40* | 39 | *14* | 49 | *52* |
| 10 | *63* | 20 | *68* | 30 | *11* | 40 | *73* | 50 | *30* |

- 8 Next carefully transfer your results from the Outcomes Sheet to the Summary Sheet, putting a tally mark for each card for a given Olympian.
- Then convert the tally marks for each Olympian to a digit (0, 1, 2, 3, 4, etc.).
- You should get a Summary Sheet something like the one below.

- **Famous Olympians: Summary Sheet**
- (Step 8 completed)
- Surname: Forename: Class:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | \| **1** | 11 | \| **1** | 21 | \| **1** | 31 | | 41 | \| **1** |
| 2 | | 12 | | 22 | \|\| **2** | 32 | | 42 | \| **1** |
| 3 | | 13 | \|\| **2** | 23 | | 33 | | 43 | |
| 4 | | 14 | \|\|\| **3** | 24 | | 34 | \|\| **2** | 44 | |
| 5 | | 15 | \| **1** | 25 | | 35 | | 45 | \| **1** |
| 6 | | 16 | \| **1** | 26 | \| **1** | 36 | | 46 | |
| 7 | | 17 | | 27 | \|\| **2** | 37 | \| **1** | 47 | |
| 8 | | 18 | | 28 | | 38 | | 48 | |
| 9 | \| **1** | 19 | | 29 | \|\| **2** | 39 | \| **1** | 49 | |
| 10 | | 20 | \| **1** | 30 | \| **1** | 40 | \| **1** | 50 | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 51 | | 61 | | 71 | | 81 | | 91 | ||| 3 |
| 52 | | 1 | 62 | | 72 | | 82 | | 92 | |
| 53 | | 63 | | 1 | 73 | | 1 | 83 | || 2 | 93 | |
| 54 | | 1 | 64 | | 74 | | 84 | | 94 | |
| 55 | | 65 | | 75 | | 85 | | 95 | || 2 |
| 56 | | 66 | | 1 | 76 | | 1 | 86 | | 96 | |
| 57 | | 1 | 67 | | 77 | | 1 | 87 | | 97 | | 1 |
| 58 | | 68 | | 1 | 78 | | 1 | 88 | | 98 | |
| 59 | | 1 | 69 | | 79 | | 89 | | 99 | || 2 |
| 60 | | 70 | | 80 | | 90 | | 1 | 100 | |

- 9 Then count the number of zeros and record the result in the Summary table at the bottom of the Summary Sheet.
- Do the same for the number of ones, twos, threes, etc.
- You should finish up with a table of results something like (but probably not the same as) the one below.

| Cards | Olympians | Cards | Olympians |
|---|---|---|---|
| 0 | ***62*** | 5 | ***0*** |
| 1 | ***28*** | 6 | ***0*** |
| 2 | ***8*** | 7 | ***0*** |
| 3 | ***2*** | 8 | ***0*** |
| 4 | ***0*** | 9 | ***0*** |

- **10 Calculation for the average number of 2-carders in a pack of 50 cards, ie the calculation for E(N = 2)**
- a) Consider a particular Olympian, say Olympian 36.
- **One way** in which Olympian 36 can be a 2-carder is if **only cards 1 and 2, and no others**, are for Olympian 36.
- The probability of this is = $(0.01)^2 \times (0.99)^{48} \approx 0.0000617$
- b) But {1,2} is just one of the 2-member subsets of { 1,2,3,…,50}
- The number of 2-member subsets of {1,2,3,…,50} is $^{50}C_2$
- and $^{50}C_2 = (50\times49)\div(2\times1) = 1225$
- So, the probability of Olympian 36 being a 2-carder $\approx 0.0000617 \times 1225$
- $= 0.0756$

- c) But there are 100 Famous Olympians, so the average number of 2-carder Olympians $\approx 100 \times 0.0756 = 7.56 \approx 7.6$
- i.e the Expected Number of 2-carders in a pack of 50 cards is 7.6
- 11 The following spreadsheet shows the results of the simulations from 25 students.
- The average for the number of 2-carders is 8.0, which is close to the theoretical result of 7.6

- Simulation results from a class of 25 students

| 2019 | Sim 1 | | Cards | | | | Total | Total | |
|---|---|---|---|---|---|---|---|---|---|
| Initials | 0 | 1 | 2 | 3 | 4 | ≥ 5 | Cards | Olympians | |
| AM | 60 | 31 | 8 | 1 | | | 50 | 100 | 1 |
| AL | 59 | 32 | 9 | | | | 50 | 100 | 1 |
| CD | 63 | 25 | 11 | 1 | | | 50 | 100 | 1 |
| GP | 64 | 24 | 10 | 2 | | | 50 | 100 | 1 |
| GA | 60 | 31 | 8 | 1 | | | 50 | 100 | 1 |
| HE2 | 61 | 30 | 7 | 2 | | | 50 | 100 | 1 |
| HE | 58 | 35 | 6 | 1 | | | 50 | 100 | 1 |
| HI2 | 61 | 29 | 9 | 1 | | | 50 | 100 | 1 |
| PW | 60 | 31 | 8 | 1 | | | 50 | 100 | 1 |
| PH | 59 | 33 | 7 | 1 | | | 50 | 100 | 1 |
| PH2 | 60 | 31 | 8 | 1 | | | 50 | 100 | 1 |
| SG | 63 | 26 | 7 | 3 | | | 49 | 99 | 1 |
| SG2 | 61 | 30 | 7 | 2 | | | 50 | 100 | 1 |

| 2019 | Sim 1 | | Cards | | | | Total | Total | |
|---|---|---|---|---|---|---|---|---|---|
| Initials | 0 | 1 | 2 | 3 | 4 | ≥ 5 | Cards | Olympians | |
| SS | 61 | 29 | 9 | 1 | | | 50 | 100 | 1 |
| TC | 62 | 26 | 12 | | | | 50 | 100 | 1 |
| TC2 | 60 | 31 | 8 | 1 | | | 50 | 100 | 1 |
| TA | 60 | 31 | 8 | 1 | | | 50 | 100 | 1 |
| WJ | 61 | 29 | 9 | 1 | | | 50 | 100 | 1 |
| WR | 61 | 30 | 7 | 2 | | | 50 | 100 | 1 |
| XN | 60 | 34 | 5 | 2 | | | 50 | 101 | 1 |
| YJ | 60 | 32 | 6 | 2 | | | 50 | 100 | 1 |
| YJ2 | 62 | 28 | 8 | 2 | | | 50 | 100 | 1 |
| NN1 | 62 | 29 | 7 | 1 | 1 | | 50 | 100 | 1 |
| NN2 | 63 | 24 | 12 | 1 | | | 51 | 100 | 1 |
| NN3 | 59 | 34 | 5 | 2 | | | 50 | 100 | 1 |
| **Av:** | **60.8** | **29.8** | **8.0** | **1.3** | **0.0** | **0.0** | | | **25** |

- 12 The question is now, “What use can be made of this (and other similar examples)?
- You might consider trying to influence schools in your region/state to include Monte Carlo simulation as a topic in their mathematics programs.
- If you don’t try, it is unlikely that the existing situation will change.


- **Dr Nathan (Norm) Hoffman**
- **Mathematical Association of Western Australia**

- **November, 2019**