Example Context Elaboration: Cereal Boxes

Focus: Simulations

Achievement objective S7-4

In a range of meaningful contexts, students will be engaged in thinking mathematically and statistically. They will solve problems and model situations that require them to:

Investigate situations that involve elements of chance:

A comparing theoretical continuous distributions, such as the normal distribution, with experimental distributions

B calculating probabilities, using such tools as two-way tables, tree diagrams, simulations, and technology

Cereal Boxes

Introduction: A cereal company is running a promotion. Each packet of cereal contains one of three animal cards.

Problem

Sunita and Henry are investigating probability situations which cannot easily be solved theoretically. They are trying to answer the question: What is the typical number of boxes of cereal that they would need to buy to get a full set of animal cards?

Plan

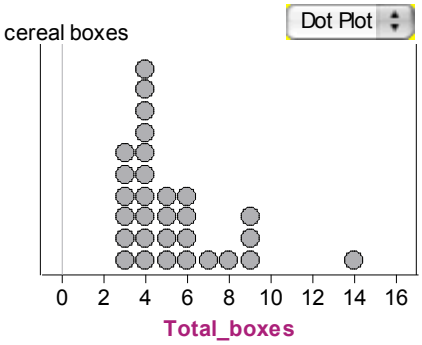
They simulate the situation for a set of three cards by using the random number function on their calculator with the formula 3×ran#+1 (to simulate buying a box with card #1, 2 or 3) and reading the integer values until they have obtained a one, a two and a three, repeating 30 times and recording the results in a table.

Data

trial	roll 1	roll 2	roll 3	roll 4	roll 5	roll 6	roll 7	roll 8	roll 9	roll 10	roll 11	roll 12	roll 13	roll 14	Total boxes
1	3	3	3	2	3	1									6
2	1	3	2												3
3	2	1	3												3
4	3	1	3	1	1	3	2								7
5	1	1	3	3	1	2									6
6	2	2	2	3	1										5
7	2	1	3												3
8	2	1	3												3
9	2	1	1	3											4
10	3	1	1	2											4
11	1	2	1	3											4
12	2	2	1	3											4
13	1	2	1	2	2	3									6
14	2	3	2	3	1										5
15	1	1	2	3											4
16	3	2	3	2	2	2	3	2	1						9
17	3	2	3	3	3	1									6
18	2	1	3	2											4
19	3	1	2												3
20	1	2	3												3
21	3	2	3	3	3	2	2	2	2	3	2	3	2	1	14
22	2	1	1	2	3										5
23	1	2	2	3											4
24	2	3	3	3	3	2	2	3	1						9
25	3	2	3	1											4
26	3	1	3	1	1	3	1	1	2						9
27	1	3	1	3	2										5
28	2	3	2	1											4
29	1	3	1	2											4
30	2	2	2	2	1	1	1	3							8

Analysis

They make a dot plot of the distribution of total boxes they would have bought in the trials of their simulation.



They describe what they notice about the distribution of their simulated cereal boxes. Half the time, the complete set was obtained in 3 or 4 purchases, but in one trial it took 14 boxes to get the full set of 3 cards.

They compare their dot plot to those of their classmates and notice that all are skewed to the right but that there is considerable difference in the maximum number of rolls to get a full set.

They note that the mean number of purchases to get a set of 3 in their sample was 5.3 boxes, that the median was 4 boxes, and that the IQR was 3 to 6 boxes. Recognising that the mean is likely to be affected by a few cases where many boxes will need to be purchased, they decide to use the median as the typical number of boxes one would have to buy to get a full set of three cards.

Conclusion

They believe that the sample is likely to representative of the population in general, but not of the occasional trial which would require a large number of rolls to obtain a full set. They realise that they would require a much larger sample to get a representative idea of the frequency of those unusual occasions.

Plan

Sunita and Henry then use the computer programme http://www.mste.uiuc.edu/reese/cereal/cereal.html to simulate the cereal box situation for obtaining sets of 3, 4, 5, 6, 7, and 8 cards.

Conclusion

They use the information to write a report to cereal consumers, informing them of the number of boxes they could expect to buy in order to obtain a full set of cards. Since the programme uses the mean, they use that to describe and compare the typical number of boxes expected for a full set, but include in their conclusion that the median would have been more appropriate when they write their report.

Reflection

How do simulations help us to answer chance questions when the theoretical probability is not available or difficult to work out?

Extension activity