Question 1

In the notes, one of the examples is about dice, and sets up a box model for the process of rolling a pair of fair six sided dice and summing the spots. In the notes we make 2 draws at random with replacement from the following box, and sum the draws:

$$\fbox{ \fbox{1} \fbox{2} \fbox{3} \fbox{4} \fbox{5} \fbox{6}}$$

Why not make one draw from the following box instead? (The tickets are the possible sums when we roll a pair of six sided dice.)

$$\fbox{ \fbox{2} \fbox{3} \fbox{4} \fbox{5} \fbox{6} \fbox{7} \fbox{8} \fbox{9} \fbox{10} \fbox{11} \fbox{12}}$$

If we did indeed want to model the result of two die rolls with a box from which we would only draw once, what would the box be? (List all the tickets in the box, which numbers, and how many of each.)

Solution

The second box doesn't work because the tickets won't be equally likely if it is to represent the outcome space of rolling a pair of dice. The correct box to use would be one that reflects the probabilities of the outcomes. For example the chance of the sum being 2 is 1 in 36, but the chance of the sum being 7 is 6 in 36. The tickets in the box have to reflect these unequal chances. The correct box, from which we can draw one ticket representing the outcome of the sum of spots of a pair of dice would be: $$\fbox{ 1 \fbox{2} 2 \fbox{3} 3 \fbox{4} 4 \fbox{5} 5 \fbox{6} 6 \fbox{7} 5 \fbox{8} 4 \fbox{9} 3 \fbox{10} 2 \fbox{11} 1 \fbox{12}}$$

Question 2

If we are setting up a box for modeling the number of heads in three tosses of a fair coin, would either of the boxes below work? If not, why not?

(a) $\fbox{\fbox{0} \fbox{1}}$ - Draw three times at random with replacement, and sum the draws.
(b) $\fbox{\fbox{0} \fbox{1} \fbox{2} \fbox{3}}$ - Draw once, the result is the number of heads.

Solution

The second one does not work, for the same reason as the first one above.

Question 3

A standard deck of cards has 52 cards, consisting of 4 "suits" (hearts , diamonds , spades ♠, and clubs ♣). Each suit has 13 cards (Ace, King, Queen, Jack, 2, 3, 4, 5, 6, 7, 8, 9, and 10). Half of the cards are red (hearts and diamonds) and half of the cards are black (spades and clubs). I would like to shuffle the cards, draw **five** cards with replacement, and count the number of **hearts** in these five cards. What would be the box model for this set up?

Question 4

One ticket will be drawn at random from each of the two boxes below:

$A: \fbox{\fbox{1} \fbox{2} \fbox{3}}$

$B: \fbox{ \fbox{1} \fbox{2} \fbox{3} \fbox{4}}$

(a) What is the chance the number drawn from A is greater than the one drawn from B?

(b) What is the chance that the number drawn from A is equal to the one drawn from B?

(c) What is the chance the number drawn from A is smaller than the one drawn from B?

Question 5

I want to estimate the proportion of people in Berkeley who speak at least two languages. I stand at the corner of University and Shattuck and ask each person who goes by how many languages they speak, and keep a count of how many speak at least two. Can I set this up as a box model? If so, how? If not, why not?

Question 6

The company Chegg wants to justify its business model and decides to conduct an opinion poll. It sends out a survey to its 7.8 million subscribers (as of 2021) asking if they have ever cheated on an assignment using Chegg. It gets back 1.5 million responses of which 87% of the respondents say that they have never used Chegg to cheat on an assignment.

(a) Could this survey be set up as a box model? If so, what tickets would go in the box and how many would we draw? If not, why not?

(b) Is this sample of 1.5 million a representative sample of high school and college students?