1. Box with 400 zeros and 600 ones. 1000 draws with replacement. Which best describes the situation?
   a. Exactly 600 ones will be drawn
   b. Almost certainly, exactly 600 ones will be drawn.
   c. Most likely, the number of ones will be slightly different than 600.

The computation: Let the random variable \( X \) = number of ones drawn.

\[
\text{Prob (} X = 600) = \text{choose}(1000, 600) \times 600^{0.6} \times 400^{0.4}
\]

by binomial formula.

\[
\text{dbinom}(600,1000, 0.6) = 0.025745 = \text{less than a three one percent chance of exactly 600 ones.}
\]

NOTE: \text{choose} is the R command for \( \frac{\text{factorial}(1000)}{\text{factorial}(600) \times \text{factorial}(400)} \)

\text{dbinom} is the binomial density function, which will give the probability of exactly 600 ones
in 1000 draws with a 0.6 chance of success on any draw

To get a sense that the exact expected outcome is rare, scale the problem down to 10 draws:

\[
\text{Prob (} X = 6) = \text{choose}(10,6) \times 6^6 \times 4^4 = 210 \times 0.0467 \times 0.0256 = 0.2508;
\]

the most likely single outcome, but has a chance of about 25 percent. Some other outcome is more likely
than the exact expected value.

2. If the 10,000 draws are without replacement, you will get the entire population, so exactly 6000 ones will be
drawn.

3. If you've lost 10 times in a row, the "law of averages" predicts nothing about what will happen on the next
spin of a fair roulette wheel. Both the gambler who thinks he is "due for a win" and the bystander who thinks
that his "luck is cold" are wrong.

* 4. Sixty or Six Hundred? Die to be rolled \( N \) times; you win $1 if one of the following conditions is met. For
each case, would you prefer \( N \) to be 60 rolls or 600 rolls? Note that the probability of an ace on any roll is \( 1/6 = 0.1667 \) or 16.67%.

   a. win if ace shows more than 20 percent of the time. \( N = 60 \) is better; the law of averages does
      say that the percentage of deviations from EV of 16.67% will become smaller.
   b. win if ace shows more than 15 percent of the time. \( N = 600 \) is better.
   c. win if ace shows between 15 and 20 percent of the time. \( N = 600 \) is better.
   d. win if the percentage of aces is exactly 16.67%. Chances are poor either way, but \( N = 60 \) offers some
      hope. See prob. 1; also compute:

\[
\text{dbinom}(10, 60, 1/6)) = 0.1370 = 13.7 \text{ percent.}
\]

\[
\text{dbinom}(100, 600, 1/6) = 0.0437 = \text{4.37 percent}
\]

As the number of possible outcomes becomes greater, the chance of exactly the expected
outcome declines.

5. If coin is tossed 100 times, the only way to get a percentage of 50 percent heads is to get exactly 50 heads!
The text problem is a garbled statement of the law of averages: the deviation from 50 percent heads
is likely to be smaller with 100 tosses than with 10 tosses.

* 6. Binomial Distribution. Assuming 2 child families and even odds of male and female children, the chance
of both being of the same sex will be 50 percent. (Calculate all binomial probabilities, and let \( X \) = number of
male children;

\[
\text{Pr (} X = 0) = \text{Pr (F and F)} = 0.5 \times 0.5 = .25
\]

\[
\text{Pr (} X = 1) = \text{Pr [ (F and M) OR (M and F)] = 0.5 \times 0.5 + 0.5 \times 0.5 = .25 + .25 = 0.5}
\]

\[
\text{Pr (} X = 2) = \text{Pr (M and M) = 0.5 \times 0.5 = .25}
\]

Note that \( X = 0 \) OR \( X = 2 \) meets the condition that both are of the same sex.
A deviation from this percentage is more likely if you look at a smaller number of families.
* 7. Multiple guess test. (25 questions, 5 options each, 0.2 chance of correct answer, 4 points per question, and a one point deduction for wrong answer. There should be 5 tickets in the box, with 25 draws. Only one ticket has the winning score of 4; the remainder have a score of -1 on them.  
   The mean of the box is 0, and the EV of sheer guessing is 25 * 0 = 0.  
   The SD of the box is 2; so the variance is 4. For the sum of 25 draws, the variance is 25 * 4 = 100, and hence the standard error of the exam score will be 10.
   
   Note that negative exam scores are possible due to the penalty, if there are more than 20 mistakes. Use Binomial(25, .2) to get the probability of any specific number of questions right; you should find there is a 42.07% percent chance of getting a negative score (X <= 4) by guessing.

* 8. Roulette box model. Gambler plays 50 turns of the roulette wheel, betting on 4 numbers; giving 4/38 chance of winning on any one turn. A win means a win of $8; a loss means a loss of $1. The box model should model the payoffs with 38 tickets, of which just 4 are labeled $8, and the other 34 labeled minus $1. The net gain in 50 plays will be the sum of 50 draws from that box.

9. Box with more red marbles than blue ones. You win $10,000 if at the end of N draws (with replacement), a red marble is drawn more often than a blue one. Would you prefer 100 draws or 200 draws? This is much harder to show definitively: see note 6 on page A17 for the proof. But it is easy to simulate for a given probability of drawing a red marble, which must be at least a bit more than .5 (say .51).

   The probability of getting LESS than half red will be given by
   \[ \text{pbinom(N/2, N, 0.51)} \]
   If N = 100 draws, \( \text{pbinom(50, 100, 0.51)} = 0.4599 \), so there is a .5401 chance of more red marbles
   If N = 200 draws, \( \text{pbinom(100, 200, 0.51)} = 0.4158 \), so there is a .5842 chance of more red marbles.

   Repeat this with a probability of 0.6 of drawing a red marble to see that the larger number of draws is even more to be preferred if there are many more red marbles than blue ones.

* 10. Sums and averages. Consider 200 draws made with replacement from the box [-3 -2 -1 0 1 2 3 ]
   a. If sum of the 200 draws is 30, what is the average? Answer: 30/200 = +.15
   b. If sum of the 200 draws is -20, the average is -20 / 200 = -0.10
   c. In general, average = Sum of draws / Number of draws.
   d. Consider the alternatives:
      i. Win $1 million if sum of 200 draws is between -5 and +5
      ii. Win $1 million if average is between -0.025 and +0.025
   
   The two alternatives have the same chance, since 5 / 200 = 0.025.

   NOTE WELL: once we have the SUMS of draws, we will also have the AVERAGES = Sums / Number of draws AND can find the percentages.

   If the sums were from a box with zeros and ones only -- say that in problem 10 we were interested in the chance of losing money, so replaced every number less than zero with "1", resulting in the box [ 1 1 1 0 0 0 0 ] -- the AVERAGE number of losses = Sum / number of draws would also be the PERCENTAGE of losses.

   This chapter applies not only to SUMS, but also to PERCENTAGES and to AVERAGES. Chapter 21 will return to the problem of the accuracy of PERCENTAGES, and chapter 23 will deal with the accuracy of PERCENTAGES. Hence understanding this problem will give you a strong basis for understanding those chapters.