Chapter 11 – Understanding Randomness

1. **Coin toss.**

   A coin flip is random, because the outcome cannot be predicted beforehand.

2. **Casino.**

   If the outcome of the video roulette game cannot be predicted ahead of time, then it is truly random. Each of the individual outcomes (numbers 01-36, plus 0 and 00 for American roulette or 01-36, plus 0, for European roulette) should be equally likely. Probably, this is not technically the case, since the video roulette machine would use a random number generator. These generators are only pseudorandom.

3. **The lottery.**

   In state lotteries, a machine pops up numbered balls. If the lottery were truly random, the outcome could not be predicted and the outcomes would be equally likely. It is random only if the balls generate numbers in equal frequencies.

4. **Games**

   Answers may vary.

   Rolling one or two dice: If the dice are fair, then each outcome, 1 through 6 should be equally likely.

   Spinning a spinner: Each outcome should be equally likely, but the spinner might be more likely to land on one outcome than another due to friction or design.

   Shuffling cards and dealing a hand: If the cards are shuffled adequately (7 times for riffle shuffling), the cards will be approximately equally likely to be in any given hand.

5. **Birth defects.**

   Answers may vary. Generate two-digit random numbers, 00-99. Let 00-02 represent a defect. Let 03-99 represent no defect.

6. **Colorblind.**

   Answers may vary. Generate random digits 0-9. Let 0 represent colorblind. Let 1-9 represent no color perception defect.

7. **Geography.**

   a) Looking at pairs of digits, the first state number is 45, Vermont. The next set is ignored since there is no 92nd state. The next state number is 10, Georgia.

   b) Continuing along, the next state number is 17, Kentucky. The next state number, 10, is ignored, since Georgia was already assigned. The final state number is 22, Michigan.

8. **Get rich.**

   Looking at pairs of digits, you would choose 43, ignore 68, since it is not a possible lottery pick, choose 09, ignore 87, choose 50, choose 13, ignore 09, since you already chose that number, choose 27. Your numbers are 43, 9, 50, 13, 27.

9. Play the lottery.

If the lottery is random, it doesn't matter if you play the same favorite "lucky" numbers or if you play different numbers each time. All numbers are equally likely (or, rather, UNLIKELY) to win.

10. Play it again, Sam.

If the lottery is random, it doesn't matter if you play random numbers or not. All numbers are equally likely (or, rather, UNLIKELY) to win.

11. Bad simulations

a) The outcomes are not equally likely. For example, the probability of getting 5 heads in 9 tosses is not the same as the probability of getting 0 heads, but the simulation assumes they are equally likely.

b) The even-odd assignment assumes that the player is equally likely to score or miss the shot. In reality, the likelihood of making the shot depends on the player's skill.

c) Suppose a hand has four aces. This might be represented by 1,1,1,1, and any other number. The likelihood of the first ace in the hand is not the same as for the second or third or fourth. But with this simulation, the likelihood is the same for each.

12. More bad simulations.

a) The numbers would represent the sums, but the sums are not all equally likely. For example, the probability of rolling a 7 is 6/36, but the probability of getting a 2 is only 1/36. The simulation assumes they are equally likely.

b) The number of boys in a family of 5 children is not equally likely. For example, having a total of 5 boys is less likely than having 3 boys out of 5 children. The simulation assigns the same likelihood to each event.

c) The likelihood for out, single, double, triple, and home run are not the same. The outcome of an at bat depends on the player's skill. The simulation assumes that these outcomes are equally likely.

13. Wrong conclusion.

The conclusion should indicate that the simulation **suggests** that the average length of the line would be 3.2 people. Future results might not match the simulated results exactly.

14. Another wrong conclusion.

The simulation **suggests** that 24% of the people might contract the disease. The simulation does not represent what happened, but what might have happened.

15. Election.

a) Answers will vary. A component is one voter voting. An outcome is a vote for our candidate. Using two random digits, 00-99, let 01-55 represent a vote for your candidate, and let 55-99 and 00 represent a vote for the underdog.

b) A trial is 100 votes. Examine 100 two-digit random numbers and count how many simulated votes are cast for each candidate. Whoever gets the majority of the votes wins the trial.

c) The response variable is whether the underdog wins or not. To calculate the experimental probability, divide the number of trials in which the simulated underdog wins by the total number of trials.

16. Two pair, or three of a Kind?

a) Answers will vary. A component is picking a single card. An outcome is the suit and denomination of the card. To simulate picking a card, generate two random digits from 00-99, let 01-52 represent the respective cards in the deck. Ignore 53-99 and 00. Alternatively, you could generate a random digit 0-9, and let 1 = spades, 2 = clubs, 3 = hearts, and 4 = diamonds. Ignore 5-9 and 0. Then generate a two digit random number 00-99, representing the denomination (01 = ace, 02 = two,…, 11 = Jack, 12 = Queen, 13 = King), ignoring 14-99 and 00.

b) A trial is a single 5-card hand. Use five sets of random numbers, ignoring repeated cards. If you were actually drawing cards, you couldn't have more than one of each card in your hand.

c) The response variable is whether the simulated hand had Two Pair, Three of a Kind, or neither. To find the experimental probability of any event, divide the number of occurrences of that event by the total number of trials.

17. Cereal.

Answers will vary. A component is the simulation of the picture in one box of cereal. One possible way to model this component is to generate random digits 0-9. Let 0 and 1 represent Tiger Woods, 2-4 represent David Beckham, and 5-9 represent Serena Williams. Each trial will consist of 5 random digits, and the response variable will be whether or not a complete set of pictures is simulated. Trials in which at least one of each picture is simulated will be a success. The total number of successes divided by the total number of trials will be the simulated probability of ending up with a complete set of pictures. According to the simulation, the probability of getting a complete set of pictures is expected to be about 51.5%.

18. Cereal, again.

Answers will vary. A component is the simulation of the picture in one box of cereal. One possible way to model this component is to generate random digits 0-9. Let 0 and 1 represent Tiger Woods, 2-4 represent David Beckham, and 5-9 represent Serena Williams. Each trial will consist of generating random numbers until a 0 or 1 is generated. The response variable will be the number of digits generated until the first 0 or 1. The total number of digits generated divided by the total number of trials will be the simulated average number of boxes required to get a Tiger Woods picture. According to the simulation, in order to be reasonably assured of getting a Tiger Woods picture, expect to buy about 5 boxes.

19. Multiple choice

Answers will vary. A component is one multiple-choice question. One possible way to model this component is to generate random digits 0-9. Let digits 0-7 represent a correct answer, and let digits 8 and 9 represent an incorrect answer. Each trial will consist of 6 random digits. The response variable is whether or not all 6 simulated questions are answered correctly (all 6 digits are 0-7). The total number of successes divided by the total number of trials will be the simulated probability of getting all 6 questions right. According to the simulation, the probability of getting all 6 multiple-choice questions correct is expected to be about 26%.

20. Lucky guessing?

Answers will vary. A component is one multiple-choice question. One possible way to model this component is to generate random digits 0-9. Let the digit 0 represent a correct answer, and let digits 1, 2, and 3 represent an incorrect answer. Ignore digits 4-9. Each trial will consist of 6 usable random digits. The response variable is whether or not all 6 simulated questions are answered correctly. The total number of successes divided by the total number of trials will be the simulated probability of getting all 6 questions right. Few simulations will have any trials getting all 6 correct, leading us to conclude that the probability of getting all 6 questions correct is very small. (The true probability is 0.00024). It isn't likely that your friend is telling the truth.

21. Beat the lottery.

a) Answers based on your simulation will vary, but you should win about 10% of the time.

b) You should win at the same rate with any number.

22. Random is as random does.

Answers based on your simulation will vary, but you should win about 10% of the time. Playing randomly selected lottery numbers offers no advantage to picking your own.

23. It evens out in the end.

Answers based on your simulation will vary, but you should win about 10% of the time. Playing lottery numbers that have turned up the least in recent lottery drawers offers no advantage. Each new drawing is independent of recent drawings.

24. Play the winner?

Answers based on your simulation will vary, but you should win about 10% of the time. Playing lottery numbers that have won in recent lottery drawers offers no advantage. Each new drawing is independent of recent drawings.

25. Driving test.

Answers will vary. A component is one drivers test, but this component will be modeled differently, depending on whether or not it is the first test taken. One possible way to model this component is to generate pairs of random digits 00-99. Let 01-34 represent passing the first test and let 35-99 and 00 represent failing the first test. Let 01-72 represent passing a retest, and let 73-99 and 00 represent failing a retest. To simulate one trial, generate pairs of random numbers until a pair is generated that represents passing a test. Begin each trial using the "first test" representation, and switch to the "retest" representation if failure is indicated on the first simulated test. The response variable is the number of simulated tests required to achieve the first passing test. The total number of simulated tests taken divided by the total number of trials is the simulated average number of tests required to pass. According to the simulation, the number of driving tests required to pass is expected to be about 1.9.

26. Still learning?

Answers will vary. A component is one drivers test, but this component will be modeled differently, depending on whether or not it is the first test taken. One possible way to model this component would be to generate pairs of random digits 00-99. Let 01-34 represent passing the first test and let 35-99 and 00 represent failing the first test. Let 01-72 represent passing a retest, and let 73-99 and 00 represent failing a retest. To simulate one trial, generate pairs of random numbers until a pair is generated that represents passing a test. Begin each trial using the "first test" representation, and switch to the "retest" representation if failure is indicated on the first simulated test. The response variable is whether or not the drivers test is passed within two attempts. The total number of simulated ***failed*** tests divided by the total number of trials is the simulated percentage of those tested who do not have a driver's license after two attempts. According to the simulation, the percentage that still do not pass within 2 tests is expected to be about 17%.

27. Basketball strategy.

Answers will vary. A component is one foul shot. One way to model this component would be to generate pairs of random digits 00-99. Let 01-72 represent a made shot, and let 73-99 and 00 represent a missed shot. The response variable is the number of shots made in a "one and one" situation. If the first shot simulated represents a made shot, simulate a second shot. If the first shot simulated represents a miss, the trial is over. The simulated average number of points is the total number of simulated points divided by the number of trials. According to the simulation, the player is expected to score about 1.24 points.

28. Blood donors.

Answers will vary. A component is one donor. One possible way to model this component is to generate pairs of random digits 00-99. Let 01-44 represent a type O donor, and let 45-99 and 00 represent a donor who is not type O. The response variable is the number of pairs of digits generated until 3 type O donors are simulated. Once 3 type O donors are simulated, the trial is over. The simulated average number of donors required is the total number of pairs of digits generated divided by the total number of trials. According to the simulation, about 6.8 donors are required to be reasonably assured of getting 3 type O donors.

29. Free groceries.

Answers will vary. A component is the selection of one card with the prize indicated. One possible way to model the prize is to generate pairs of random digits 00-99. Let 01-10 represent $200, let 11-20 represent $100, let 21-40 represent $50, and let 41-99 and 00 represent $20. Repeated pairs of digits must be ignored. (For this reason, a simulation in which random digits 0-9 are generated with 0 representing $200, 1 representing $100, etc., is NOT acceptable. Each card must be individually represented.) A trial continues until the total simulated prize is greater than $500. The response variable is the number of simulated customers until the payoff is greater than $500. The simulated average number of customers is the total number of simulated customers divided by the number of trials. According to the simulation, about 10.2 winners are expected each week.

30. Find the ace.

Answers will vary. A component is turning over one card. One way to model the cards turned over is to generate random digits 0-9. Let the digit 0 represent the ace, and let digits 1, 2, 3, and 4 each represent one of the other four cards. Ignore digits 5-9. A trial consists of simulating turning over the cards until the ace is drawn. Each card must be represented individually; repeated digits must be ignored. The response variable is the number of simulated cards drawn until the ace is drawn, with $100 being awarded if the ace is drawn first, and $50, $20, $10, or $5 if the ace is drawn second, third, fourth, or fifth, respectively. The simulated average dollar amount of music the store is expected to give away is the total dollar amount of music given away divided by the number of trials. According to the simulation, the dollar amount given away is expected to be about $37.

31. The family.

Answers will vary. Each child is a component. One way to model the component is to generate random digits 0-9. Let 0-4 represent a boy and let 5-9 represent a girl. A trial consists of generating random digits until a child of each gender is simulated. The response variable is the number of children simulated until this happens. The simulated average family size is the number of digits generated in each trial divided by the total number of trials. According to the simulation, the expected number of children in the family is about 3.

32. A bigger family.

Answers will vary. Each child is a component. One way to model the component is to generate random digits 0-9. Let 0-4 represent a boy and let 5-9 represent a girl. A trial consists of generating random digits until two children of each gender are simulated. The response variable is the number of children simulated until this happens. The simulated average family size is the number of digits generated in each trial divided by the total number of trials. According to the simulation, the expected number of children in the family is slightly less than 6.

33. Dice game.

Answers will very. Each roll of the die is a component. One way of modeling this component is to generate random digits 0-9. The digits 1-6 correspond to the numbers on the faces of the die, and digits 7-9 and 0 are ignored. A trial consists of generating random numbers until the sum of the numbers is exactly 10. If the sum exceeds 10, the last roll must be ignored and simulated again, but still counted as a roll. The response variable is the number of rolls until the sum is exactly 10. The simulated average number of rolls until this happens is the total number of rolls simulated divided by the number of trials. According to the simulation, expect to roll the die about 7.5 times.

34. Parcheesi.

Answers will vary. Each roll of two dice is a component. One way of modeling this component is to generate random digits 0-9. The digits 1-6 correspond to the numbers on the faces of the die, and digits 7-9 and 0 are ignored. For this simulation, look at the digits in usable pairs of digits, and consider the sum, as well as the numbers themselves. A trial consists of generating usable pairs of digits until the sum is 3, or until at least one of the dice shows a 3. The response variable is the number of pairs of usable numbers generated until this happens. The simulated average number of rolls is the total number of rolls divided by the number of trials. According to the simulation, expect to roll the dice about 2.7 times.

35. The hot hand.

Answers may vary. Each shot is a component. One way of modeling this component is to generate pairs of random digits 00-99. Let 01-65 represent a made shot, and let 66-99 and 00 represent a missed shot. A trial consists of 20 simulated shots. The response variable is whether or not the 20 simulated shots contained a run of 6 or more made shots. To find the simulated percentage of games in which the player is expected to have a run of 6 or more made shots, divide the total number of successes by the total number of trials. According to the simulation, the player is expected to make 6 or more shots in a row in about 40% of games. This isn't unusual. The announcer was wrong to characterize her performance as extraordinary.

36. The World Series.

Answers may vary. Each game is a component. One way of modeling this component is to generate pairs of random digits 00-99. Let 01-55 represent a win by the favored team, and let 56-99 and 00 represent a win by the underdog. A trial consists of generating pairs until one team has 4 simulated wins. The response variable is whether or not the underdog wins. The simulated percentage of World Series wins is the total number of successes divided by the total number of trials. According to simulation, the underdog is expected to win the World Series about 39% of the time.

37. Teammates.

Answers will vary. Each player chosen is a component. One way to model this component is to generate random numbers 0-9. Let 1 and 2 represent the first couple, 3 and 4 the second couple, 5 and 6 the third couple, and 7 and 8 the fourth couple. Ignore 9 and 0. A trial consists of generating random digits, ignoring repeats, and organizing them into pairs, until pairs representing the first three teams are generated. (The final team is assigned by default.) The response variable is whether or not each of the simulated teams is a pairing other than 1-2, 3-4, 5-6, or 7-8. The simulated percentage of the time this is expected to happen is the total number of successes (times that the pairings are *different* than the couples) divided by the total number of trials. According to the simulation, all players are expected to be paired with someone other than the person with whom he or she came to the party about 37.5% of the time.

38. Second team.

Answers will vary. Each player chosen is a component. One way to model this component is to generate random numbers 0-9. Let digits 1-4 represent the four players who are to be chosen. Ignore digits 5-9 and 0. A trial consists of generating a sequence of random numbers that represents the order in which the cards where chosen. Since each number represents a person, and people cannot be chosen more than once, ignore repeated numbers. The response variable is whether or not any digit in the generated sequence matches the corresponding digit in the sequence 1234 (or any other sequence of the four numbers, as long as it is determined ahead of time). The simulated percentage of the time this is expected to happen is the total number of successes (times that the sequences have no matching corresponding digits) divided by the number of trials. According to the simulation, all players are expected to be paired with someone other than the person with whom he or she came to the party about 37.5% of the time.

39. Job discrimination?

Answers may vary. Each person hired is a component. One way of modeling this component is to generate pairs of random digits 00-99. Let 01-10 represent each of the 10 women, and let 11-22 represent each of the 12 men. Ignore 23-99 and 00. A trial consists of 3 usable pairs of numbers. Ignore repeated pairs of digits, since the same man or woman cannot be hired more than once. The response variable is whether or not all 3 simulated hires are women. The simulated percentage of the time that 3 women are expected to be hired is the number of successes divided by the number of trials. According to the simulation, the 3 people hired will all be women about 7.8% of time. This seems a bit strange, but not quite strange enough to be evidence of job discrimination.

40. Cell phones.

Answers will vary. Each driver is a component. One way to model this component is to generate random digits 00-99. Let 01-12 represent a driver that is talking on his or her cell phone, and let 13-99 and 00 represent a driver that is not talking on his or her cell phone. A trial consists of 10 pairs of digits. The response variable is whether or not at least 4 of the simulated drivers were talking on their cell phones. The simulated percentage of the time that 4 or more drivers of ten are talking on their cell phones if the true rate of usage is 12%is the number of successes divided by the total number of trials. You should expect to find 4 or more drivers talking among 10 drivers only about 2% of the time. Based on what you saw at the bus stop, you'd suspect that the legislator's claim of 12% usage is probably too low.