FAT CHANCE: PROBABILITY FROM THE GROUND UP

Professional Skills 3

Auteur
Jeffrey Stynen
04/01/2025

Contents

Introduction

In this course[1], I learned the basics of probability and statistics, with the necessary background in combinatoric. This gives me a better understanding of key metrics when analysing data, which will elevate my ability to draw insights from complex datasets. The course got pretty abstract at times, but with the very helpful videos from the professors, I was able to manage and understand all the topics discussed below. In the summary below I will always go trough the theory first and then explain some of the most important exercises from every theme. This way, I will have a solid foundation to fall back on when I ever need to look something up related to probability or statistics.

1 Basic Counting

The first chapter lays the foundation for understanding probability and statistics by exploring the essential principles of counting. These principles, while seemingly simple, are critical building blocks for more advanced topics. By developing a clear grasp of how to systematically count objects, sequences, and subsets, learners gain the tools necessary to approach real-world problems involving combinatorics and probability. This chapter introduces key concepts such as inclusive counting, the multiplication principle, factorials, and the subtraction principle, each accompanied by practical examples to illustrate their application.

Theory

The first chapter of the course focuses on the essential yet powerful concepts of counting. While the act of counting may appear straightforward, understanding its mathematical foundation is critical for developing a solid grasp of probability and statistics. This chapter begins with a simple exercise: counting the number of integers between two values. By examining how many integers lie between k and n, learners are introduced to the inclusive counting formula, given as $n - k + 1$. For instance, the number of integers between 3 and 7 is calculated as $7 - 3 + 1 = 5$, covering the values 3, 4, 5, 6, and 7. This straightforward calculation sets the stage for more advanced methods.

Next, the course introduces the multiplication principle, a cornerstone of combinatorics. This principle states that if one event can occur in n_1 ways and a second independent event can occur in n_2 ways, the total number of possible outcomes is $n_1 \times n_2$. More generally, if there are k events, each with $n_1, n_2, \ldots, n_k$ outcomes, the total number of outcomes is calculated as:

$$n_1 \times n_2 \times n_3 \times \cdots \times n_k$$

In a simple example, if you have 4 shirts and 3 pairs of pants to choose from, the total number of combinations of outfits is $4 \times 3 = 12$. The multiplication principle is then extended to sequences. When repetition is allowed, the number of sequences of k objects chosen from a collection of n objects is given by n^k. For example, forming a three-digit PIN using the digits 0 through 9 results in $10^3 = 1000$ possible PINs. However, when repetition is not allowed, the calculation becomes slightly more complex. In this case, the number of sequences is determined by multiplying n choices for the first object, $n-1$ for the second, $n-2$ for the third, and so on, continuing until k choices are made. The formula for this is:

$$n \times (n-1) \times (n-2) \times \cdots \times (n-k+1)$$

For example, arranging three books from a shelf of five results in $5 \times 4 \times 3 = 60$ possible arrangements. At this stage, the course avoids introducing factorials, keeping the focus on understanding the mechanics of this formula.

Later in the chapter, the concept of factorials is briefly hinted at as a way to simplify calculations involving repeated multiplication. Factorials, denoted as $n!$, represent the product of all positive integers up to n, expressed as:

$$n! = n \times (n-1) \times (n-2) \times \cdots \times 1$$

While factorials are not formally introduced in this chapter, they are presented as a preview of concepts to come, particularly in permutations and combinations.

Finally, the chapter explores the subtraction principle, a method for simplifying problems involving subsets. This principle relies on the idea of exclusion: to count the desired subset, count the total and subtract the cases that do not meet the criteria. For example, if a group of 20 students includes 5 who failed an exam, the number of students who passed can be calculated as $20 - 5 = 15$. This principle is particularly useful in problems where directly counting the desired elements is challenging or impractical.

Overall, this chapter emphasizes the importance of mastering basic counting principles. From the simple act of counting integers to the powerful multiplication and subtraction principles, these foundational tools prepare learners to tackle more advanced topics in probability and statistics. The progression from straightforward examples to increasingly complex scenarios ensures that learners build confidence and develop a clear understanding of these essential concepts.

Exercise: Phone Numbers

A phone number has seven digits and cannot begin with a 0. How many phone numbers contain the sequence 123?

This problem can be solved by applying the multiplication and subtraction principles.

First, we determine the total placements of the sequence 123. The sequence "123" must appear as a block within the 7-digit phone number. This means it can start at positions 1, 2, 3, 4, or 5 (inclusive). For each placement:

If "123" starts in position 1, the remaining four digits can be any number from 0 to 9. This results in:

$$10^4$$

If "123" starts in position 2, the digit before the block (in position 1) cannot be 0, and the three digits after the block can be any digit. This results in:

$$9 \times 10^3$$

Similarly, if "123" starts in positions 3, 4, or 5, each placement allows:

$$9 \times 10^3$$

Adding all these cases together gives:

$$10^4 + 4 \times 9 \times 10^3$$

Next, we account for overlaps or invalid cases. Specifically, if the sequence "123" starts at position 1 and the number starts with a 0, this violates the rule that a phone number cannot begin with a 0. Such overlaps are carefully analysed and found to total 29 invalid arrangements. Subtracting these from the total gives the final count:

$$10^4 + 36 \times 10^3 - 29$$

The total number of valid phone numbers containing the sequence "123" is:

$$10^4 + 36 \times 10^3 - 29$$

2 Advanced Counting

Building on the foundational principles introduced in Chapter 1, this chapter explores more advanced counting techniques. These concepts extend the idea of systematically counting objects to more complex scenarios, including arrangements and collections involving multiple categories. By the end of this chapter, learners will have mastered key tools such as the binomial coefficient, multinomial coefficient, and methods for calculating collections with or without repetition.

Theory

The chapter begins with the introduction of the binomial coefficient, a powerful tool for counting the number of ways to choose k objects from a set of n objects without repetition. The binomial coefficient is denoted as:

$$\binom{n}{k}$$

and is calculated using the formula:

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

For example, if you want to choose 3 objects from a set of 5, the number of possible combinations is:

$$\binom{5}{3} = \frac{5!}{3!\,(5-3)!} = \frac{120}{6 \times 2} = 10$$

This tool is particularly useful for problems involving collections, where the order of selection does not matter.

Once the binomial coefficient is introduced, the course demonstrates its applications. These include counting subsets, solving combinatorial problems, and understanding how it relates to the expansion of binomials (as seen in the binomial theorem). For example, the coefficient of $x^k y^{n-k}$ in the expansion of $(x+y)^n$ is given by:

$$\binom{n}{k}$$

This section helps solidify the concept by connecting it to real-world and algebraic examples.

After mastering the binomial coefficient, learners are introduced to the multinomial coefficient, which generalizes the binomial coefficient to scenarios involving multiple categories. The multinomial coefficient counts the number of ways to divide n objects into k groups of specified sizes $n_1, n_2, \ldots, n_k$. It is denoted as:

$$\binom{n}{n_1, n_2, \ldots, n_k}$$

and is calculated using the formula:

$$\binom{n}{n_1, n_2, \ldots, n_k} = \frac{n!}{n_1!\, n_2! \, \ldots \, n_k!}$$

For example, if you have 10 objects and want to divide them into 3 groups of sizes 4, 3, and 3, the number of ways to do this is:

$$\binom{10}{4, 3, 3} = \frac{10!}{4!\, 3!\, 3!}$$

This tool is essential for problems involving more complex divisions and classifications.

The chapter concludes by addressing the final missing piece in the counting table: calculating collections with repetition. This involves determining the number of ways to choose k objects from a set of n objects when repetition is allowed. The solution to this problem is given by the formula:

$$\binom{n+k-1}{k}$$

This formula arises from the "stars and bars" theorem, which visualizes the problem as dividing k indistinguishable objects among n distinguishable categories.

For example, if you want to distribute 5 identical candies among 3 children, the number of ways to do so is:

$$\binom{5+3-1}{5} = \binom{7}{5} = 21$$

This final piece completes the framework for counting, enabling learners to handle any combination of sequences and collections, with or without repetition. Expanding our toolkit by introducing advanced counting techniques, we are now able to tackle more complex problems. The binomial coefficient provides a method for counting collections without repetition, while the multinomial coefficient extends this idea to multiple categories. Finally, the method for counting collections with repetition completes the framework, ensuring that learners can confidently approach a wide variety of combinatorial problems. These tools are foundational for understanding probability, combinatorics, and applications in fields ranging from algebra to data analysis.

Exercise 1: Justifying the Binomial Coefficient Formula

Prove the recursive formula for the binomial coefficient:

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

This formula can be understood by considering a set of n objects and the process of choosing k objects from this set. To derive the formula, we divide the set of n objects into two parts: one specific object, which we will call A, and the remaining $n-1$ objects.

There are two possible cases for forming a subset of size k:

1. In the first case, the object A is not included in the subset. This means that all k objects must be chosen from the remaining $n-1$ objects. The number of ways to do this is represented by $\binom{n-1}{k}$.

2. In the second case, the object A is included in the subset. Since A occupies one of the k spots, we are left with $k-1$ spots to be filled by choosing $k-1$ objects from the remaining $n-1$ objects. The number of ways to do this is represented by $\binom{n-1}{k-1}$.

By combining these two cases, we see that the total number of ways to choose k objects from a set of n objects is the sum of the two possibilities. This leads directly to the formula:

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}.$$

This recursive property is a fundamental result in combinatorics and forms the basis for constructing Pascal's Triangle and other combinatorial tools.

Exercise 2: Identity from the Binomial Theorem

What identity do we get when we substitute the values $a = 1$ and $b = -1$ into the binomial theorem?

The binomial theorem states that for any positive integer n, we can expand $(a+b)^n$ as:

$$(a+b)^n = \binom{n}{0}a^n b^0 + \binom{n}{1}a^{n-1}b^1 + \binom{n}{2}a^{n-2}b^2 + \cdots + \binom{n}{n}a^0 b^n.$$

To solve this exercise, we substitute $a = 1$ and $b = -1$ into the formula. Substituting these values gives:

$$(1-1)^n = \binom{n}{0}(1)^n(-1)^0 + \binom{n}{1}(1)^{n-1}(-1)^1 + \binom{n}{2}(1)^{n-2}(-1)^2 + \cdots + \binom{n}{n}(1)^0(-1)^n.$$

On the left-hand side, $(1-1)^n = 0^n$, which is equal to 0 for any positive integer n. On the right-hand side, the powers of 1 simplify to 1, and the equation becomes:

$$0 = \binom{n}{0}(-1)^0 + \binom{n}{1}(-1)^1 + \binom{n}{2}(-1)^2 + \cdots + \binom{n}{n}(-1)^n.$$

The alternating signs come from the powers of -1, with the first term being positive, the second term negative, the third term positive, and so on. In other words, the sum of the terms alternates in sign:

$$\sum_{k=0}^{n} \binom{n}{k} (-1)^k = 0.$$

This result shows that the alternating sum of binomial coefficients equals zero for any positive integer n.

Exercise 3: Distributing Sandwiches Equally

It's a beautiful spring day, so you and your 14 siblings have decided to go on a picnic. Your uncle, who owns a deli, has provided sandwiches: five ham and cheese, five turkey, and five egg salad. How many ways are there of distributing the sandwiches to everyone, including yourself? Choose the best answer.

This problem involves distributing 15 sandwiches into three distinct groups, where each group contains exactly 5 sandwiches of the same type. Since the sandwiches of the same type are indistinguishable (e.g., all five ham and cheese sandwiches are identical), we need to calculate the number of unique arrangements of these three groups.

The formula for this type of arrangement is based on the multinomial coefficient, which is given by:

$$\frac{n!}{n_1!\, n_2!\, n_3!}.$$

Here:

- n is the total number of sandwiches, which is 15.
- n_1, n_2, and n_3 are the sizes of the three groups, which are 5, 5, and 5 respectively.

Substituting these values into the formula, we have:

$$\frac{15!}{5! \times 5! \times 5!}.$$

This calculation gives the total number of ways to arrange the 15 sandwiches into three distinct groups where each group contains exactly 5 identical sandwiches.

The total number of ways to distribute the sandwiches is:

$$\frac{15!}{5! \times 5! \times 5!}.$$

3 Basic Probability

Probability is the mathematical study of uncertainty, allowing us to quantify the likelihood of different outcomes in everyday situations. From predicting the result of a coin flip to calculating the odds of a rare poker hand, probability provides a framework for analysing randomness and making informed decisions.

This chapter builds upon the counting techniques introduced in previous sections and applies them to solve probability problems. We begin with fundamental concepts, such as defining probability in terms of favourable and total outcomes, before exploring more complex cases involving dice rolls and card games. As the chapter progresses, we introduce key principles, including the multiplication and subtraction rules, which enable us to compute probabilities in increasingly sophisticated scenarios. By the end, readers will be well-equipped to tackle real-world problems involving randomness and uncertainty.

Theory

This section introduces the concept of probability and its application in real-world scenarios. Probability is defined as the ratio of favourable outcomes to the total number of possible outcomes. Mathematically, this is expressed as:

$$P = \frac{k}{n},$$

where k represents the number of favourable outcomes and n represents the total number of outcomes. The section begins with simple examples, such as flipping coins, and progresses to more complex scenarios like rolling dice and calculating probabilities for poker hands.

The first example, flipping a coin, illustrates the basics of probability. With two possible outcomes—heads or tails—the total number of outcomes is $n = 2$. For a single flip, the probability of either heads or tails is $P = \frac{1}{2}$. When multiple coins are flipped, previously learned counting principles, such as the multiplication principle, are applied to calculate the probabilities of specific sequences of outcomes.

The section then moves on to rolling dice, incorporating more advanced counting techniques. For instance, when rolling two six-sided dice, the total number of outcomes is calculated as $6 \times 6 = 36$. Probabilities for specific scenarios, such as rolling doubles or achieving a sum of 7, are determined by identifying favourable outcomes and dividing by the total number of outcomes.

The lesson then transitions to a more complex application: poker hands. A standard deck of cards consists of 52 cards, divided into four suits—spades, hearts, diamonds, and clubs—each containing 13 denominations (ace, 2 through 10, jack, queen, and king). By the multiplication principle, the total number of cards is calculated as $4 \times 13 = 52$. A poker hand is defined as a collection of five cards drawn from the deck, where order does not matter and no card is repeated. The total number of possible poker hands is calculated using the binomial coefficient:

$$\binom{52}{5} = \frac{52!}{5!(52-5)!} = 2,598,960.$$

The probabilities of being dealt specific types of poker hands are then explored. For example, a four of a kind hand consists of four cards of the same denomination and one additional card. The number of such hands is calculated as:

$$13 \times 48 = 624,$$

where 13 represents the number of possible denominations and 48 represents the remaining cards for the fifth card. The probability of being dealt a four of a kind is:

$$P = \frac{624}{2,598,960} \approx 0.00024,$$

which is about 1 in 4,000. Another example is the full house, which consists of three cards of one denomination and two cards of another denomination. The number of such hands is calculated using the multiplication principle:

$$13 \times \binom{4}{3} \times 12 \times \binom{4}{2} = 13 \times 4 \times 12 \times 6 = 3,744.$$

The probability of being dealt a full house is:

$$P = \frac{3,744}{2,598,960} \approx 0.00144,$$

or about 1 in 700. As expected, four of a kind hands are rarer than full houses, reflecting their higher rank in poker.

The section also examines three of a kind, which includes three cards of the same denomination and two other cards of different denominations. To ensure the calculation excludes hands that qualify as full houses, the subtraction principle is applied. The number of hands that meet the three-of-a-kind criteria is:

$$54,912 = 58,656 - 3,744,$$

where 58,656 is the total number of hands with three cards of the same denomination, and 3,744 is the number of full house hands. The probability of being dealt a three of a kind is:

$$P = \frac{54,912}{2,598,960} \approx 0.0211,$$

or about 1 in 47.

For learners seeking an additional challenge, the optional section on bridge hands explores the distribution of cards in the game of bridge. This topic builds on the multinomial coefficient to calculate probabilities for specific distributions, providing a deeper application of combinatorics.

In conclusion, the Basic Probability section demonstrates how counting principles can be applied to calculate probabilities in increasingly complex scenarios. Beginning with simple examples like flipping coins and rolling dice, the section progresses to calculating poker probabilities using tools like the binomial coefficient, the multiplication principle, and the subtraction principle. These tools provide a strong foundation for solving real-world probability problems and serve as essential building blocks for more advanced topics.

Exercise 1: Probability of Coin Flips

Suppose you flip a fair coin 10 times. What is the probability of getting between four and six heads in total?

This exercise involves calculating the probability of achieving a total of either four, five, or six heads out of 10 coin flips. Since the coin is fair, the probability of each outcome (heads or tails) is equal, and we can apply the binomial distribution formula to determine the probability for each case.

The probability of getting exactly k heads in n flips of a fair coin is given by:

$$P_k = \frac{\binom{n}{k}}{2^n},$$

Where: $\binom{n}{k}$ is the binomial coefficient, n is the total number of flips, and k is the desired number of heads.

Here, $n = 10$ and we are interested in the cases $k = 4$, $k = 5$, and $k = 6$. The probability for each case is calculated as follows:

1. For $k = 4$:
$$P_4 = \frac{\binom{10}{4}}{2^{10}},$$
2. For $k = 5$:
$$P_5 = \frac{\binom{10}{5}}{2^{10}},$$
3. For $k = 6$:
$$P_6 = \frac{\binom{10}{6}}{2^{10}}.$$

The total probability of getting between four and six heads is the sum of these individual probabilities:

$$P_{\text{total}} = P_4 + P_5 + P_6 = \frac{\binom{10}{4}}{2^{10}} + \frac{\binom{10}{5}}{2^{10}} + \frac{\binom{10}{6}}{2^{10}}.$$

Factoring out the common denominator, this simplifies to:

$$P_{\text{total}} = \frac{\binom{10}{4} + \binom{10}{5} + \binom{10}{6}}{2^{10}}.$$

The probability of getting between four and six heads in 10 flips of a fair coin is:

$$\frac{\binom{10}{5} + 2\binom{10}{4}}{2^{10}}.$$

Exercise 2: Probability of a Busted Flush

What is the probability of being dealt a busted flush: four cards of the same suit, but a fifth card of a different suit?

This exercise requires calculating the number of hands where four cards are of the same suit, but the fifth card is of a different suit. To solve this, we apply the multiplication principle and the concept of combinations.

There are 4 suits in a standard deck, so there are 4 choices. Each suit has 13 cards, and we need to choose 4 of them. This can be done in: $\binom{13}{4}$ ways.

The fifth card must be from a different suit. Since there are 3 remaining suits, each with 13 cards, there are: $3 \times 13 = 39$ ways to choose this card.

The total number of busted flush hands is therefore:

$$4 \times \binom{13}{4} \times 39.$$

The total number of poker hands is given by:

$$\binom{52}{5}.$$

Thus, the probability of being dealt a busted flush is:

$$P = \frac{4 \times \binom{13}{4} \times 39}{\binom{52}{5}}.$$

After simplifying, the probability of being dealt a busted flush is:

$$\frac{12 \times \binom{13}{4} \times 13}{\binom{52}{5}}.$$

4 Expected Value

Expected value is one of the most powerful concepts in probability and statistics, providing a way to quantify long-term outcomes in uncertain situations. Whether analysing gambling strategies, evaluating investment opportunities, or making decisions in everyday life, expected value serves as a crucial tool for assessing risks and rewards.

In this chapter, we explore the fundamental principles of expected value and how it can be used to make informed decisions. We begin by defining expected value as a weighted average of all possible outcomes, where each outcome is multiplied by its corresponding probability. Through practical examples—such as rolling dice and comparing different betting strategies—we illustrate how expected value helps to predict long-term trends, even when individual outcomes vary significantly. Additionally, we introduce techniques like the subtraction principle to simplify probability calculations in complex scenarios.

By the end of this chapter, readers will not only understand how to compute expected value but also how to apply it to real-world decision-making, ensuring a more strategic approach to uncertainty.

Theory

The section on Expected Value introduces a fundamental concept in probability and statistics: calculating the long-term average outcome of a random process. This concept is crucial for analysing games, investments, and other situations involving uncertainty. The general formula for the expected value of a random variable X is given as:

$$E(X) = \sum_{i=1}^{n} P_i \cdot X_i,$$

where X_i represents each possible outcome, P_i is the probability of that outcome, and n is the total number of possible outcomes. In essence, the expected value is a weighted average of all possible outcomes, with the weights being the probabilities of each outcome. A critical reminder is that all probabilities must add up to 1:

$$\sum_{i=1}^{n} P_i = 1.$$

This is illustrated with examples from dice games. For instance, consider a game where you roll a six-sided die and receive a payout equal to the number rolled in dollars. The possible outcomes are $X_i = 1, 2, 3, 4, 5, 6$, and each outcome has a probability of $P_i = \frac{1}{6}$. The expected value is calculated as:

$$E(X) = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6.$$

Simplifying further, this becomes:

$$E(X) = \frac{1}{6} \cdot (1 + 2 + 3 + 4 + 5 + 6) = \frac{1}{6} \cdot 21 = 3.5.$$

This means that, on average, a player can expect to win \$3.50 per roll over many repetitions of the game. While any individual roll may differ, the expected value represents the long-term average outcome.

The subtraction principle is also introduced as a helpful tool for calculating probabilities in scenarios where directly counting favourable outcomes is complex. The principle involves subtracting the number of unfavourable outcomes from the total number of possible outcomes. For example, if you want to calculate the probability of rolling at least a 4 on a six-sided die, it's often simpler to subtract the probability of rolling less than 4 from 1:

$$P(\text{at least 4}) = 1 - P(\text{less than 4}).$$

This method becomes especially useful in games with multiple outcomes, as it reduces the need to count all individual possibilities explicitly.

The section concludes by showing how expected value can be applied to make strategic decisions in games and other uncertain situations. For example, consider a choice between two bets: one offering a guaranteed payout of \$4, and another involving a dice game where rolling a 6 pays \$20 while rolling any other number pays \$0. For the first bet, the expected value is simply $E(X) = 4$. For the second bet, the expected value is calculated as:

$$E(X) = \frac{1}{6} \cdot 20 + \frac{5}{6} \cdot 0 = \frac{20}{6} = 3.33.$$

Since the expected value of the first bet (4) is higher than that of the second bet (3.33), the first bet is the better choice in the long run. This example highlights the power of expected value as a decision-making tool, helping individuals quantify and compare the long-term benefits of different options.

The key takeaway from this section is that expected value provides a systematic way to evaluate random processes and make informed decisions. By combining tools like the subtraction principle and the general formula for expected value, learners gain a deeper understanding of how to navigate uncertainty strategically.

Exercise 1: Probability of Being Paid Nothing in a Dice Game

Suppose you play a game by rolling seven dice. If you roll five of a kind, the payoff is \$50. If you get six of a kind, the payoff is \$500. If you get all seven of a kind, the payoff is \$5,000. What is the probability that you'll be paid nothing at all on the roll?

To solve this problem, we calculate the probability of the complementary event: the probability of getting at least five of a kind. Once this is determined, the probability of being paid nothing is simply the complement of this value, calculated as $1 - P(\text{at least five of a kind})$.

For all seven dice to show the same number, there are six possible outcomes (one for each number on the die). The probability of rolling seven of a kind is:

$$\frac{1}{6^6}.$$

To roll six of a kind, six dice must show the same number, while the seventh die can be any other number. There are six ways to choose the repeated number, and five options for the seventh die. The number of ways to arrange the seven dice is given by:

$$6^5 \times 7,$$

and the probability is:

$$\frac{5 \times 7}{6^6}.$$

For five of a kind, five dice must show the same number, and the other two dice must show two different numbers. This can be calculated as:

$$\frac{5^2 \times \binom{7}{5}}{6^6}.$$

The total probability of rolling at least five of a kind is the sum of these probabilities:

$$\frac{5^2 \cdot \binom{7}{5}}{6^6} + \frac{5 \cdot 7}{6^6} + \frac{1}{6^6}.$$

Finally, the probability of being paid nothing is the complement of this value:

$$1 - \left(\frac{5^2 \cdot \binom{7}{5}}{6^6} + \frac{5 \cdot 7}{6^6} + \frac{1}{6^6} \right).$$

The probability of being paid nothing is:

$$1 - \left(\frac{5^2 \cdot \binom{7}{5}}{6^6} + \frac{5 \cdot 7}{6^6} + \frac{1}{6^6} \right).$$

Exercise 2: Comparing Two Games

Here are two variations of a one-card game:

- Game 1: You are dealt one card from a standard deck of 52 cards. You win \$10 if you get a face card, and otherwise, the payoff is twice the number on your card. For example, \$2 for an Ace, \$4 for a 2, and so on, up to \$20 for a 10.

- Game 2: You are also dealt one card from the same deck. You win \$20 for a face card. If the card is a number card (Ace counts as 1), you get \$10 if the number is even and \$5 if the number is odd.

Which game is better?

To determine which game is better, we calculate the expected value for each game.

Game 1: The deck consists of 52 cards, of which:

- 12 are face cards (J, Q, K) with a payout of \$10 each.
- 40 are number cards (Aces through 10), with payouts calculated as twice the card's value.

The expected value is:

$$E(\text{Game 1}) = \frac{12}{52}\cdot 10+\frac{4}{52}\cdot 2+\frac{4}{52}\cdot 4+\frac{4}{52}\cdot 6+\frac{4}{52}\cdot 8+\frac{4}{52}\cdot 10+\frac{4}{52}\cdot 12+\frac{4}{52}\cdot 14+\frac{4}{52}\cdot 16+\frac{4}{52}\cdot 18+\frac{4}{52}\cdot 20.$$

Simplify:

$$E(\text{Game 1}) = \frac{12\cdot 10 + (4\cdot 2 + 4\cdot 4 + 4\cdot 6 + 4\cdot 8 + 4\cdot 10 + 4\cdot 12 + 4\cdot 14 + 4\cdot 16 + 4\cdot 18 + 4\cdot 20)}{52}.$$

Calculating the numerator:

$$E(\text{Game 1}) = \frac{120 + (8 + 16 + 24 + 32 + 40 + 48 + 56 + 64 + 72 + 80)}{52}.$$

$$E(\text{Game 1}) = \frac{120 + 440}{52} = \frac{560}{52} \approx 10.77.$$

Game 2: The deck consists of:

- 12 face cards with a payout of \$20.
- 40 number cards (Aces through 10), split into:
- 20 even cards (2, 4, 6, 8, 10) with a payout of \$10 each.
- 20 odd cards (Aces, 3, 5, 7, 9) with a payout of \$5 each.

The expected value is:

$$E(\text{Game 2}) = \frac{12}{52}\cdot 20 + \frac{20}{52}\cdot 10 + \frac{20}{52}\cdot 5.$$

Simplify:

$$E(\text{Game 2}) = \frac{240 + 200 + 100}{52} = \frac{540}{52} \approx 10.38.$$

Comparison: The expected value of Game 1 is approximately \$10.77, while the expected value of Game 2 is approximately \$10.38. Therefore, **Game 1** has a higher expected value and is the better choice.

5 Conditional Probability

Probability often deals with uncertainty, but what happens when new information becomes available? Conditional probability allows us to refine our predictions by incorporating additional knowledge, making it an essential tool in decision-making, statistics, and real-world problem-solving.

This chapter explores how probabilities change when new evidence is introduced. We begin with the famous Monty Hall problem, a counter-intuitive game show puzzle that highlights the power of conditional probability in strategic decision-making. Through this example, we see how updating probabilities based on new information can dramatically alter outcomes.

Next, we introduce the concept of correlation, which measures the strength of relationships between variables. While correlation can reveal patterns in data, we emphasize the critical distinction between correlation and causation, illustrating how apparent connections can sometimes be misleading.

The chapter then progresses to Bayes' Theorem, a fundamental principle for updating probabilities when new evidence is observed. This theorem has wide-ranging applications, from medical diagnostics to spam filtering and artificial intelligence. Through practical examples—such as evaluating the accuracy of a medical test—we demonstrate how Bayes' Theorem helps refine predictions based on prior knowledge and observed data.

By the end of this chapter, learners will have a solid understanding of conditional probability, its real-world applications, and how to use it to make better-informed decisions in uncertain situations.

Theory

The chapter on Conditional Probability explores how the likelihood of an event changes when additional information is available. This concept is crucial for understanding real-world scenarios where events are interdependent. The chapter begins with the famous Monty Hall problem, which provides an intuitive example of how conditional probability works.

The Monty Hall problem involves a game show where a contestant chooses one of three doors. Behind one door is a car (the prize), and behind the other two are goats. After the contestant makes their choice, the host, Monty Hall, opens one of the remaining doors, revealing a goat. The contestant is then given the option to stick with their original choice or switch to the other unopened door. Conditional probability is used to analyse whether switching or staying is the better strategy.

To understand the Monty Hall problem, we calculate the probability of winning the car when the contestant switches doors. Using the law of total probability, the probability of winning (denoted as $P(W)$) can be expressed as:

$$P(W) = P(A) \cdot P(W \text{ assuming } A) + P(B) \cdot P(W \text{ assuming } B),$$

where:

$P(A)$ is the probability the car is behind the door the contestant initially chose,

$P(B)$ is the probability the car is behind one of the other doors,

$P(W \text{ assuming } A)$ is the probability of winning given the car is behind the initially chosen door,

$P(W \text{ assuming } B)$ is the probability of winning given the car is behind one of the other doors.

Substituting the probabilities:

$P(A) = \frac{1}{3}$ (since there is a 1-in-3 chance the car is behind the initially chosen door),

$P(B) = \frac{2}{3}$ (since there is a 2-in-3 chance the car is behind one of the other two doors),

$P(W \text{ assuming } A) = 0$ (because switching would lose in this case),

$P(W \text{ assuming } B) = 1$ (because switching would win if the car is behind one of the other two doors).

The probability of winning by switching is:

$$P(W) = \frac{1}{3} \cdot 0 + \frac{2}{3} \cdot 1 = \frac{2}{3}.$$

In contrast, the probability of winning by sticking with the original choice is:

$$P(W \text{ — Stay}) = \frac{1}{3}.$$

This counter-intuitive result demonstrates that switching doors doubles the contestant's chances of winning the car. The Monty Hall problem vividly illustrates how new information (the host revealing a goat) changes the probabilities of the outcomes and highlights the importance of conditional probability in decision-making.

As the chapter progresses, the focus shifts to correlation, which measures the strength and direction of the relationship between two variables. Correlation is quantified using the correlation coefficient, typically denoted as ρ or r. Values of ρ range from -1 to 1, where:

$\rho = 1$ indicates a perfect positive correlation (as one variable increases, the other increases proportionally),

$\rho = -1$ indicates a perfect negative correlation (as one variable increases, the other decreases proportionally),

$\rho = 0$ indicates no linear correlation.

The distinction between correlation and causation is emphasized, with real-world examples highlighting how a strong correlation does not necessarily imply that one variable causes the other.

Towards the end of the chapter, learners are introduced to Bayes' Theorem, a powerful tool for updating probabilities based on new evidence. Bayes' Theorem is stated as:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)},$$

where $P(A|B)$ is the updated probability of event A given event B, $P(B|A)$ is the likelihood of event B occurring if A is true, $P(A)$ is the prior probability of A, and $P(B)$ is the total probability of event B. Bayes' Theorem is particularly useful in fields such as medicine, machine learning, and decision-making.

For example, consider a medical test for a rare disease where:

$P(\text{Disease}) = 0.01$ (1
$P(\text{Positive Test — Disease}) = 0.95$ (95
$P(\text{Positive Test — No Disease}) = 0.05$ (5

Using Bayes' Theorem, we can calculate the probability that a person actually has the disease given a positive test result:

$$P(\text{Disease — Positive Test}) = \frac{P(\text{Positive Test — Disease}) \cdot P(\text{Disease})}{P(\text{Positive Test})}.$$

The denominator $P(\text{Positive Test})$ can be expanded using the law of total probability:

$$P(\text{Positive Test}) = P(\text{Positive Test — Disease}) \cdot P(\text{Disease}) + P(\text{Positive Test — No Disease}) \cdot P(\text{No Disease}).$$

Substituting values and calculating, Bayes' Theorem helps quantify how much a positive test result changes the likelihood of having the disease.

In conclusion, the chapter on Conditional Probability equips learners with tools to analyse interdependent events. From understanding how new information affects probabilities (Monty Hall problem) to quantifying relationships (correlation) and updating beliefs (Bayes' Theorem), this chapter provides essential techniques for reasoning under uncertainty and making data-informed decisions.

Exercise 1: Generalized Monty Hall Problem

In Monty Hall's generalized game, there are k cars as prizes behind n doors ($1 \leq k < n-1$). The rules are the same as before: You pick a door. Monty opens a door you didn't pick to reveal a goat. Then, you have the opportunity to switch your pick. Let A denote the event of guessing a door with a car behind it, and let B denote the event of guessing a door with a goat behind it on your initial pick. Let L represent the event of losing the game if your strategy is always to switch doors. Calculate $P(L \cap B)$.

To calculate $P(L \cap B)$, we break down the problem into manageable steps.
- B: The event that the initially chosen door has a goat behind it.
- L: The event of losing when switching. This occurs if the contestant picks a goat initially, and Monty's reveal leads to switching to another goat.

The probability of picking a goat initially is:

$$P(B) = \frac{n-k}{n},$$

because there are $n-k$ doors with goats and n doors in total.
For $L \cap B$, the contestant must:
- Initially pick a goat (B).
- Switch to another goat after Monty reveals one door with a goat.

If B occurs, then there are $n-k-1$ doors remaining with goats (excluding the initial goat chosen and the one revealed by Monty). The total number of remaining doors (excluding the initial pick and Monty's reveal) is $n-2$.

Thus, the probability of switching to another goat (and losing) given that B occurred is:

$$P(L|B) = \frac{n-k-2}{n-2}.$$

Using the definition of conditional probability:

$$P(L \cap B) = P(B) \cdot P(L|B).$$

Substituting the values:

$$P(L \cap B) = \frac{n-k}{n} \cdot \frac{n-k-2}{n-2}.$$

The probability of $L \cap B$ is:

$$P(L \cap B) = \frac{(n-k)(n-k-2)}{n(n-2)}.$$

Exercise 2: Bayes Theorem

The exercise involves determining the probability that a patient is sick given a positive test result, using Bayes' Theorem. The problem provides the following information:

$P(\text{Sick}) = 0.01$ (1% of people in the population are sick),

$P(\text{Positive — Sick}) = 0.99$ (the test correctly identifies 99% of sick people),

$P(\text{Positive — Healthy}) = 0.01$ (1% false positive rate for healthy people),

$P(\text{Healthy}) = 1 - P(\text{Sick}) = 0.99$ (99% of people in the population are healthy).

We want to calculate $P(\text{Sick — Positive})$, the probability that the patient is sick given a positive test result. According to Bayes' Theorem, this is expressed as:

$$P(\text{Sick — Positive}) = \frac{P(\text{Positive — Sick}) \cdot P(\text{Sick})}{P(\text{Positive})}.$$

Here, $P(\text{Positive})$ is the total probability of testing positive, which accounts for both sick and healthy individuals. Using the law of total probability, we compute:

$$P(\text{Positive}) = P(\text{Positive — Sick}) \cdot P(\text{Sick}) + P(\text{Positive — Healthy}) \cdot P(\text{Healthy}).$$

Substituting the given values:

$$P(\text{Positive}) = 0.99 \cdot 0.01 + 0.01 \cdot 0.99 = 0.0099 + 0.0099 = 0.0198.$$

Now substituting back into Bayes' Theorem:

$$P(\text{Sick — Positive}) = \frac{0.99 \cdot 0.01}{0.0198} = \frac{0.0099}{0.0198} = 0.5.$$

Thus, the probability that the patient is sick given a positive test result is:

$$P(\text{Sick — Positive}) = \frac{1}{2}.$$

This result demonstrates that even with a highly reliable test, the low prevalence of the disease in the population significantly affects the final probability.

6 Bernoulli Trials

Many real-world processes involve repeated independent events with two possible outcomes—success or failure. From lottery tickets and coin flips to medical tests and sports series, the ability to analyze these events is crucial in probability theory. The mathematical framework for these scenarios is known as Bernoulli trials, which provide a foundation for calculating the likelihood of specific outcomes over multiple attempts.

In this chapter, we explore the Bernoulli formula, which models the probability of achieving exactly k successes in n independent trials. We begin with simple applications, such as calculating the probability of winning a lottery at least once after multiple ticket purchases. The chapter then extends to more complex scenarios, including the gambler's ruin problem, which examines long-term betting strategies and the likelihood of financial success or bankruptcy.

We also introduce the concept of absorbing barriers, which model real-world constraints such as a gambler reaching a win/loss limit or a sports team winning a championship series. By comparing different strategies—such as betting incrementally versus going all-in—we gain insights into risk management and probability-driven decision-making.

By the end of this chapter, readers will be equipped with the tools to analyze repeated independent events, assess strategic choices in games of chance, and apply probability models to practical situations in finance, sports, and beyond.

Theory

The chapter on Bernoulli Trials introduces the framework for analyzing repeated independent events with two possible outcomes: "success" and "failure." These events are modeled using the general Bernoulli formula:

$$P(k) = \binom{n}{k} \cdot p^k \cdot q^{n-k},$$

where:

$P(k)$ is the probability of exactly k successes in n trials,

p is the probability of success for a single trial,

$q = 1 - p$ is the probability of failure.

This formula is demonstrated with an example involving tickets. Suppose $p = \frac{1}{6}$ (1 in 6 tickets is a winner). The chapter raises the question: is the probability of winning better when buying 6 tickets and hoping for at least 1 win, buying 12 tickets and hoping for at least 2 wins, or buying 18 tickets and hoping for at least 3 wins? Using the Bernoulli formula, the probability of at least one success when buying 6 tickets is calculated as:

$$P(\text{at least 1 win}) = 1 - P(\text{no wins}) = 1 - \binom{6}{0} \cdot p^0 \cdot q^6 = 1 - q^6.$$

Similarly, the probabilities for buying 12 and 18 tickets can be calculated to compare strategies. This example highlights the usefulness of Bernoulli trials for analysing real-world scenarios involving repeated events.

The chapter then transitions to the gambler's ruin problem, also known as a random walk with absorbing barriers. This concept explores scenarios where a player repeatedly bets on an event (e.g., roulette) and stops playing either upon reaching a target amount or losing everything. The general formula for the probability of success in a random walk is given by:

$$x(a) = \frac{s^a - 1}{s^b - 1},$$

where:

$x(a)$ is the probability of reaching the target amount starting with a,
$s = \frac{q}{p}$ is the ratio of the probabilities of failure and success,
b is the absorbing barrier for success (e.g., the target amount),
a is the starting amount.

For example, consider a player starting with 1000, betting 1 at a time on black in roulette (48% chance of winning, $p = 0.48$, $q = 0.52$), and stopping when reaching either 2000 or 0. Using the formula, we calculate the ratio:

$$s = \frac{q}{p} = \frac{0.52}{0.48} \approx 1.0833.$$

The probability of doubling the money (starting with $a = 1000$ and stopping at $b = 2000$) is:

$$x(1000) = \frac{s^{1000} - 1}{s^{2000} - 1}.$$

The chapter compares this incremental betting strategy with betting all 1000 at once. For the all-in strategy, the probability of success is simply:

$$P(\text{all-in success}) = p = 0.48.$$

The gambler's ruin framework illustrates that while incremental betting reduces variance, it also increases the likelihood of eventual ruin due to the cumulative risk of failure.

In conclusion, the Bernoulli Trials chapter equips learners with tools to analyze repeated independent events and extends these concepts to scenarios involving random walks and absorbing barriers. These tools are particularly useful for answering questions about probabilities in gambling, investments, and other repeated decision-making processes.

Exercise 1: Tickets

In a lottery where the odds of winning are 1 in 6, we want to calculate the probability of winning exactly three times when purchasing 18 tickets. This is an example of Bernoulli trials, where we calculate the probability of exactly k successes in n independent trials. Each trial (purchasing a ticket) has two possible outcomes: success (a winning ticket) or failure (a losing ticket). The general formula for Bernoulli trials is:

$$P(k) = \binom{n}{k} \cdot p^k \cdot q^{n-k},$$

where $P(k)$ is the probability of getting exactly k successes, $\binom{n}{k}$ is the binomial coefficient (the number of ways to choose k successes out of n trials), p is the probability of success for a single trial, and $q = 1 - p$ is the probability of failure for a single trial.
For this problem:

$n = 18$ (the total number of trials or tickets purchased),
$k = 3$ (the desired number of successes),
$p = \frac{1}{6}$ (the probability of winning for a single ticket),
$q = 1 - p = \frac{5}{6}$ (the probability of losing for a single ticket).
Substituting these values into the Bernoulli formula gives:

$$P(3) = \binom{18}{3} \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^{15}.$$

We begin by calculating the binomial coefficient $\binom{18}{3}$, which represents the number of ways to choose 3 successes out of 18 trials. This is calculated as:

$$\binom{18}{3} = \frac{18!}{3!(18-3)!} = \frac{18 \cdot 17 \cdot 16}{3 \cdot 2 \cdot 1} = 816.$$

Next, substitute the binomial coefficient and the probabilities p and q into the formula:

$$P(3) = 816 \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{5}{6}\right)^{15}.$$

Now calculate the powers of p and q:
The probability $\left(\frac{1}{6}\right)^3$ is:

$$\left(\frac{1}{6}\right)^3 = \frac{1}{216}.$$

The probability $\left(\frac{5}{6}\right)^{15}$ can be left in exponential form for clarity or computed numerically if needed. Finally, substitute everything back into the formula:

$$P(3) = 816 \cdot \frac{1}{216} \cdot \left(\frac{5}{6}\right)^{15}.$$

Exercise 2: The Gamblers Ruin

In the World Series, there are two baseball teams, Team A and Team B. They play a series of games until one team has won a total of four games. The series can last for a maximum of seven games. Team A has a 40% chance of winning any given game, and Team B has a 60% chance. What are the odds of Team A winning the World Series?

This problem can be solved using the concept of conditional probabilities and binomial distributions. Since the series is best-of-seven, the possible ways Team A can win are:
-Team A wins the series in exactly four games.
-Team A wins the series in five, six, or seven games.
The probability of each of these scenarios depends on Team A winning exactly enough games to reach four victories while not allowing Team B to win four games first. To compute this, we consider the probabilities of the following outcomes:

Team A wins in exactly 4 games: For this to happen, Team A must win all four games, and the probability is:

$$P(4 \text{ wins in } 4 \text{ games}) = (0.4)^4 = 0.0256.$$

Team A wins in exactly 5 games: In this case, Team A wins 4 games and loses 1. The loss can occur in any of the first four games, and the winning game must be the fifth. The number of ways this can happen is $\binom{4}{1} = 4$, and the probability is:

$$P(4 \text{ wins in 5 games}) = \binom{4}{1} \cdot (0.4)^4 \cdot (0.6)^1 = 4 \cdot (0.4)^4 \cdot (0.6) = 4 \cdot 0.0256 \cdot 0.6 = 0.06144.$$

Team A wins in exactly 6 games: Team A wins 4 games and loses 2. The losses can occur in any of the first five games, and the winning game must be the sixth. The number of ways this can happen is $\binom{5}{2} = 10$, and the probability is:

$$P(4 \text{ wins in 6 games}) = \binom{5}{2} \cdot (0.4)^4 \cdot (0.6)^2 = 10 \cdot 0.0256 \cdot 0.36 = 0.09216.$$

Team A wins in exactly 7 games: Team A wins 4 games and loses 3. The losses can occur in any of the first six games, and the winning game must be the seventh. The number of ways this can happen is $\binom{6}{3} = 20$, and the probability is:

$$P(4 \text{ wins in 7 games}) = \binom{6}{3} \cdot (0.4)^4 \cdot (0.6)^3 = 20 \cdot 0.0256 \cdot 0.216 = 0.11059.$$

Summing up the probabilities: To find the total probability of Team A winning the series, we add up all the scenarios:

$$P(\text{Team A wins}) = P(4 \text{ wins in 4 games}) + P(4 \text{ wins in 5 games}) + P(4 \text{ wins in 6 games}) + P(4 \text{ wins in 7 games}).$$

Substituting the values:

$$P(\text{Team A wins}) = 0.0256 + 0.06144 + 0.09216 + 0.11059 = 0.16.$$

Final Answer: The probability of Team A winning the World Series is approximately 16%.

7 The Normal Distribution

Many natural and random processes exhibit patterns that, when analysed over numerous trials, reveal a distinct bell-shaped curve. This curve, known as the normal distribution, is one of the most fundamental concepts in probability and statistics. It provides a powerful model for understanding variations in measurements, outcomes of repeated experiments, and real-world phenomena ranging from test scores to financial markets.

This chapter begins with simple examples, such as flipping coins and rolling dice, to illustrate how probability distributions emerge. As the number of trials increases, the shape of the probability distribution stabilizes and begins to resemble the normal curve. This progression is formalized through the central limit theorem, which states that the sum of a large number of independent, identically distributed random variables tends to follow a normal distribution, regardless of the original distribution of the individual outcomes.

We then introduce the key characteristics of the normal distribution, including expected value (mean) and variance (spread of outcomes), which help quantify the shape and center of the distribution. These parameters allow us to describe the likelihood of outcomes within specific ranges, leading to practical applications such as quality control, risk assessment, and hypothesis testing.

To standardize comparisons across different distributions, we introduce z-scores, which measure how far an outcome deviates from the mean in terms of standard deviations. This normalization enables us to compute probabilities for various scenarios, such as determining the likelihood of rolling an unusually high total when summing 100 dice rolls.

By the end of this chapter, learners will have a deep understanding of how the normal distribution models real-world variability, how the central limit theorem explains its widespread applicability, and how tools like variance and standard deviation help measure uncertainty. These concepts form the foundation of statistical analysis and are essential for interpreting data and making informed decisions.

Theory

The chapter on normal distribution begins by exploring probability distributions through practical examples such as coin tosses and dice games. These examples illustrate how probabilities are distributed across outcomes. For instance, when flipping a fair coin multiple times, the probability of achieving a certain number of heads forms a symmetric distribution centred around the expected value. As the number of trials increases, this distribution begins to resemble a bell-shaped curve, known as the normal distribution.

The normal distribution is characterized by two key parameters: the expected value (mean) and the variance. The expected value, denoted as $ev(G)$ for a game G, is the weighted average of all possible outcomes, calculated by summing the product of each outcome's value and its probability. The variance, denoted as σ^2, measures how much the outcomes deviate from the expected value and is given by:

$$\sigma^2 = p_1(a_1 - ev(G))^2 + p_2(a_2 - ev(G))^2 + \cdots + p_n(a_n - ev(G))^2,$$

where:

p_i represents the probability of the i-th outcome,

a_i represents the value of the i-th outcome,

$ev(G)$ is the expected value of the game.

The standard deviation, σ, is the square root of the variance:

$$\sigma = \sqrt{\sigma^2}.$$

To better understand the behaviour of repeated games or trials, the chapter introduces the central limit theorem, which states that as the number of independent trials n increases, the distribution of their cumulative outcomes approaches a normal distribution, regardless of the original distribution of the individual outcomes. For repeated iterations of a game G, the expected value and variance scale linearly with the number of trials:

$$ev(G(n)) = n \cdot ev(G),$$

$$\sigma^2(G(n)) = n \cdot \sigma^2(G).$$

The chapter also introduces the concept of the normalized game, which allows us to compare games on a standardized scale by subtracting the expected value and dividing by the standard deviation. The normalized game, G_0, is expressed as:

$$G_0 = \frac{G - ev}{\sigma}.$$

The normalized game highlights how far an outcome is from the mean, measured in units of standard deviation, and is a cornerstone of standardizing data in statistics.

Through examples such as rolling dice and flipping coins, the chapter demonstrates how distributions evolve as the number of trials increases, ultimately converging to the bell-shaped curve of the normal distribution. Visualizations like bar charts of outcomes illustrate this progression, reinforcing the relationship between probabilities and the normal distribution.

In summary, the normal distribution chapter emphasizes the significance of the bell curve in understanding probability, the central limit theorem in explaining its universality, and practical tools such as variance and standard deviation to measure uncertainty in outcomes. These foundational concepts underpin much of probability and statistics, enabling a deeper understanding of patterns and randomness in real-world scenarios.

Exercise : The Normal Distribution

Say you roll 100 dice and add up the numbers showing. On average, you'd expect this sum to be 350. What are the odds it will be over 400?

To solve this, we use the properties of the normal distribution. The sum of the outcomes when rolling 100 dice follows a normal distribution because of the central limit theorem, which states that the sum of a large number of independent random variables approaches a normal distribution.

Each die roll has: an expected value, or mean, of $ev(\text{die}) = 3.5$ (the average of all possible outcomes from 1 to 6), a variance of $\sigma^2(\text{die}) = \frac{(6-1+1)^2-1}{12} = 2.9167$ (using the formula for the variance of a uniform distribution).

For 100 rolls:

$$ev(\text{sum}) = n \cdot ev(\text{die}) = 100 \cdot 3.5 = 350.$$

The variance of the sum is:

$$\sigma^2(\text{sum}) = n \cdot \sigma^2(\text{die}) = 100 \cdot 2.9167 = 291.67.$$

The standard deviation is the square root of the variance:

$$\sigma(\text{sum}) = \sqrt{291.67} \approx 17.08.$$

We are interested in the probability of the sum being greater than 400. To calculate this, we normalize 400 using the z-score formula:

$$z = \frac{X - ev(\text{sum})}{\sigma(\text{sum})},$$

where X is the target value, $ev(\text{sum})$ is the mean, and $\sigma(\text{sum})$ is the standard deviation. Substitute the values:

$$z = \frac{400 - 350}{17.08} \approx \frac{50}{17.08} \approx 2.93.$$

The z-score tells us how many standard deviations the target value (400) is above the mean (350). Using z-tables or statistical software, we find the cumulative probability corresponding to $z = 2.93$. This is approximately:

$$P(Z \leq 2.93) \approx 0.9983.$$

The probability of rolling a sum greater than 400 is the complement of this value:

$$P(X > 400) = 1 - P(Z \leq 2.93) = 1 - 0.9983 = 0.0017.$$

The probability of rolling a sum greater than 400 when rolling 100 dice is approximately 0.0017, or 0.17%.

References

[1] Fat chance: Probability from the ground up. [Online]. Available: https://learning.edx.org/course/course-v1:HarvardX+FC1x+1T2024/home