MAT220 Lecture 8 Handout

The Binomial Distribution Pt. 1

Contents

In this lecture we cover one of the final topics in our discussion of probability theory: binomial experiments. As we will see, these are particular setups in which there are two possible outcomes, namely "success" or "failure". Given that there are only two possibilities, these systems are quite easy to describe and to calculate, meaning that we can relate them to our previous material in an easy way. After our presentation of the general theory, we will discuss various examples of binomial experiments.

1 A brief review: counting and distributions

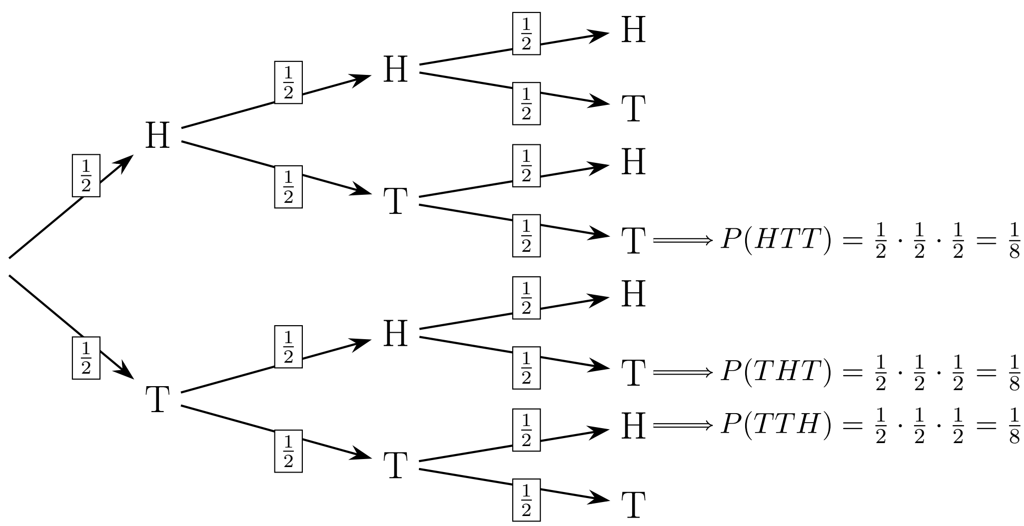
1.1 Calculating the probability of 2 Tails in 3 coin flips

Suppose that we have a fair coin, and we flip it 3 times. What is the probability of getting exactly 2 Tails? Fortunately, with a small number like 3 we can simply write out all of the possible outcomes:

- A single coin flip has 2 outcomes: H or T.
- Two coin flips have 4 outcomes: HH, HT, TH, TT.
- Three coin flips have 8 outcomes (the full sample space):

$$\{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}.$$

There are several ways we can try to calculate a probability like $P(\text{exactly 2 Ts})$. One method could be to use a decision tree (like in Lecture 7) to calculate the probability of each favourable outcome:



The three events HTT, TTH and THT are mutually exclusive, so we can simply add up these individual outcomes and this will give the probability of the compound event:

$$P(\text{exactly 2 Ts}) = P(HTT \text{ or } THT \text{ or } TTH) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}.$$

Alternatively, we could arrive at the same answer by working with a simpler calculation of compound events (like in Lecture 5). We can observe that in this situation all of the outcomes are equally likely, so we could also just use the Golden Formula directly:

$$P(\text{event}) = \frac{\text{number of outcomes favourable to the event}}{\text{total number of outcomes}}.$$

In this case there are 3 outcomes that contain precisely two tails, namely HTT, THT and TTH, so:

$$P(\text{exactly 2 Ts}) = \frac{\text{number of outcomes with 2 Ts}}{\text{total number of outcomes}} = \frac{3}{8}.$$

1.2 General solutions to this problem

If we now flip a coin n times, the sample space will have 2^n outcomes. To find the probability of "exactly r tails", we want to calculate the ratio:

$$\frac{\text{number of outcomes with exactly } r \text{ tails}}{2^n}.$$

Therefore, the problem can be solved once we find a way to count the outcomes with exactly r tails among n coin flips.

One way to figure this out is with the following: let's describe the outcome of n-many flips as a string of Hs and Ts, e.g. a string of 5 coin flips may look like: $HHTHT$. Before we actually flip the coin, we can imagine these slots as totally blank and yet to be assigned, e.g. for $n = 5$ we can picture the potential outcomes as _, _, _, _, _. Now, asking for exactly r-many tails is like asking for the ways in which we can assign r-many of these slots as Ts. Once we choose which r slots are T, every other slot is forced to be H, so each choice of Ts encodes a unique outcome. Different choices of slots give different outcomes. So, of the n available slots, we have to choose r-many of them to assign a T to. In other words, we have n many potential choices and we have to commit to exactly r-many. This is precisely a *combination*, meaning we are choosing an r-element subset of the n flip positions, and therefore the number of different choices is counted by the formula $C_{n,r}$.

In Lecture 4 we saw that the number of ways to choose r-many from a collection of size n can simply be read off of Pascal's triangle:

$$\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1
\end{array}
\qquad
\begin{array}{ccccccccccc}
 & & & & & C_{0,0} & & & & & \\
 & & & & C_{1,0} & & C_{1,1} & & & & \\
 & & & C_{2,0} & & C_{2,1} & & C_{2,2} & & & \\
 & & C_{3,0} & & C_{3,1} & & C_{3,2} & & C_{3,3} & & \\
 & C_{4,0} & & C_{4,1} & & C_{4,2} & & C_{4,3} & & C_{4,4} & \\
C_{5,0} & & C_{5,1} & & C_{5,2} & & C_{5,3} & & C_{5,4} & & C_{5,5}
\end{array}$$

Specifically, the answer can be found by going to row n (starting from row 0) and finding the entry for r (starting from $r = 0$). This will give $C_{n,r}$, which is the number we want. For instance, the entry with $r = 2$ in the third row is a 3, which corresponds to the number of ways in which we can get 2 tails from 3 coin flips.

Suppose now that we flip our coin 5 times. What is the probability of getting exactly 2 tails?

Solution

Total outcomes: $2^5 = 32$. The number of outcomes with exactly 2 tails is $C_{5,2} = 10$ (row 5 of Pascal's triangle). Therefore,

$$P(\text{exactly 2 tails}) = \frac{C_{5,2}}{2^5} = \frac{10}{32} = \frac{5}{16}.$$

1.3 Probability distributions

A variable X is called a *random variable* if the value that X takes in a given experiment or observation is determined by a random outcome. There are two main types:

- A *discrete random variable* can take on only a finite number of values, or a countable number of values.
- A *continuous random variable* can take on any of the countless number of values in an interval.

Exercise: discrete or continuous?

Classify each random variable as discrete or continuous.

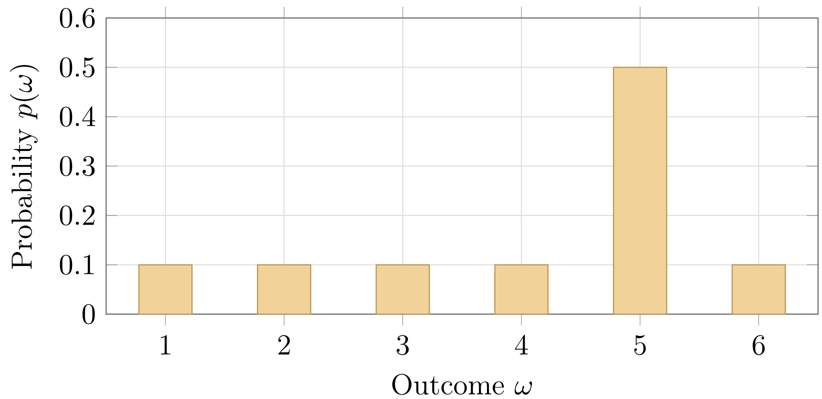
(a) The time it takes a selected student to register for the Fall term.
(b) The number of text messages received by a selected student on a randomly chosen day.
(c) The number of miles an electric vehicle can drive on a full charge.
(d) Pick a random sample of 50 registered voters in a district and find the number who voted in the last county election.

Solution

(a) Continuous.
(b) Discrete.
(c) Continuous.
(d) Discrete.

A *discrete probability distribution* is an assignment of probabilities to each distinct value of a discrete random variable. In Lecture 6, we saw that probability distributions can be visualised as bar charts. For example, the graph:

ω	$p(\omega)$
1	0.1
2	0.1
3	0.1
4	0.1
5	0.5
6	0.1



visually explains how probability is spread out amongst the possible values $1, 2, 3, 4, 5$ and 6.

Normally we write random variables with capital letters, such as X or Y. If X is allowed to take the values $1, 2, 3, 4, 5, 6$, then statements like

$$X = 4 \qquad \text{or} \qquad X = 5$$

represent possible values of the random variable. The probability that X takes the value 4 can then be written:

$$P(X = 4).$$

On a probability distribution, this is represented by the height of the bar at 4. In the example above, we have

$$P(X = 4) = 0.1.$$

We can also ask for the probability that X lies in a range of values. For example:

$$P(X < 3)$$

means the probability that X takes a value less than 3. Since the values less than 3 are 1 and 2, we can calculate the total probability by taking the sum:

$$P(X < 3) = P(X = 1) + P(X = 2) = 0.1 + 0.1 = 0.2.$$

In Lecture 6, we also saw a formula for the *expected value* of a random variable, which acted like the mean of a probability distribution. The expected value of a probabilistic system is a weighted average. For a finite discrete random variable, the expected value is

$$E(X) = \sum_{i=1}^{n} p_i x_i = p_1 x_1 + \cdots + p_n x_n.$$

This formula can be carefully derived by trying to calculate the geometric center of mass of the probability distribution – for the full explanation you should refer back to Lecture 6. Informally, $E(X)$ is the long-term expected behaviour of X.

The variance of a probability distribution measures how spread out the possible values of X are around their expected value. It is defined by

$$\mathrm{Var}(X) = E\big((X - E(X))^2\big).$$

Observe the similarities this formula has with the usual definition of population variance for a data set. In summary:

Expected value and variance

The **expected value** of a discrete random variable is

$$E(X) = \sum_{i=1}^{n} p_i x_i.$$

It is the probabilistic version of the mean, and therefore gives a measure of the center of the distribution.

The **variance** of a random variable is

$$\mathrm{Var}(X) = E\big((X - E(X))^2\big).$$

It measures how much the distribution spreads out around its expected value.

1.4 The law of large numbers

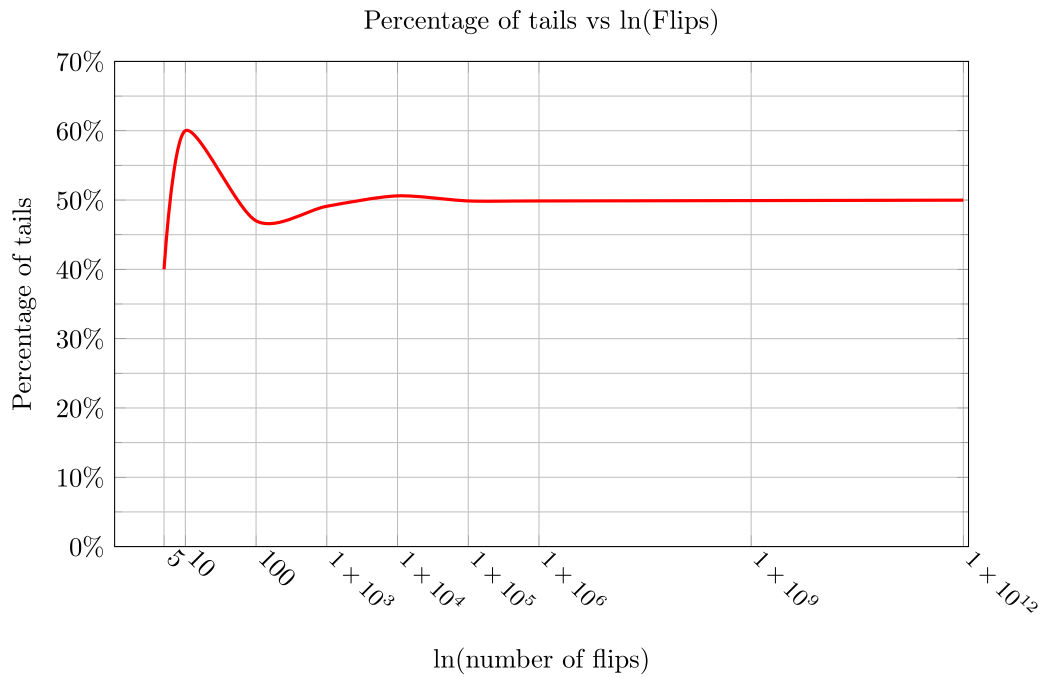
The expected value of a random variable is a way to calculate the average, long-term expected result of an experiment. We will now elaborate on this idea by introducing another interpretation of probability.

The *law of large numbers* says that if you repeat an experiment many times, then the average result tends to get closer to the theoretical average. As an example of this, suppose that we have a fair coin. We expect that the probability of flipping the coin and landing on tails is 50%, i.e. $P(T) = \frac{1}{2}$. Our interpretation of probability says that we expect there to be about 50% tails in the long run.

To see this in action, suppose that we start flipping a coin over and over again. Using a random number generator, below is an example of the proportion of tails in the sample of coin flips:

- 5 flips: HHTTH. $T = 2$, $H = 3$, so tails $= 40\%$.
- 10 flips: HHTTHTHTTT. $T = 6$, $H = 4$, so tails $= 60\%$.
- 100 flips: $T = 47$, $H = 53$, so tails $= 47\%$.
- 1000 flips: $T = 491$, $H = 509$, so tails $= 49.1\%$.
- 10000 flips: $T = 5062$, $H = 4938$, so tails $= 50.6\%$.
- 100000 flips: $T = 49872$, $H = 50128$, so tails $= 49.87\%$.
- 1 million flips: tails $= 49.87\%$.
- 1 billion flips: tails $= 49.92\%$.
- 1 trillion flips: tails $= 49.982\%$.

We can visualize this by plotting the graph (where here we take a logarithmic scale $\ln(\cdot)$ so that the coin flips fit on the same axis):



As you can see, in the beginning there is quite a lot of random fluctuation in the results. However, over time these random changes mostly cancel out, and the proportion of Tails converges to the theoretical average of 50%. This offers us one possible interpretation of probability:

The Frequentist Interpretation of probability

When we say $P(A) = x\%$, we mean that if we repeated the experiment a huge number of times, we expect the long-run proportion of outcomes corresponding to event A would get closer and closer to $x\%$.

In this sense, the expected value $E(X)$ can be better understood as the average result that we would get if we repeated the experiment over and over again under independent conditions. Eventually, the running average of this repeated activity will have results that will converge to something, and this something is the expected value. As we will see later, the expected value is especially useful when discussing Binomial experiments.

2 Example: binomial experiments

2.1 Example 1: 10 flips of a fair coin

Suppose we have a fair coin, and we flip it 10 times in a row. What is the probability of finding exactly 5 tails?

Solution

Based on the previous section, we can calculate $P(\text{exactly 5 tails})$ using the formula:

$$P(\text{exactly 5 tails}) = \frac{C_{10,5}}{2^{10}}.$$

Of course, we may use the usual formula $C_{n,r} = \frac{n!}{r!(n-r)!}$, but for narrative consistency let's instead use Pascal's triangle. The 10^{th} row of the triangle is:

$$1 \quad 10 \quad 45 \quad 120 \quad 210 \quad 252 \quad 210 \quad 120 \quad 45 \quad 10 \quad 1$$

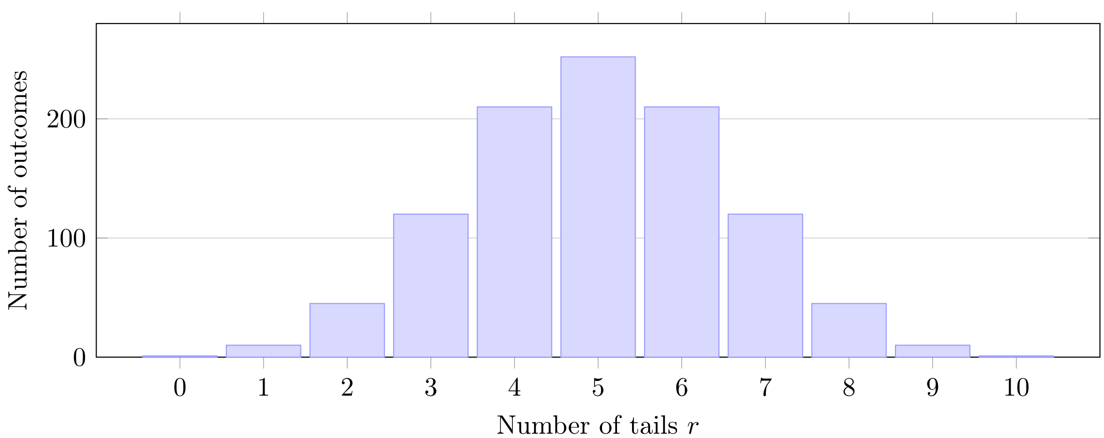
So, we move across to the entry with $r = 5$ (remembering that the triangle starts counting at 0) to see that $C_{10,5} = 252$. The total number of outcomes is $2^{10} = 1024$. Therefore:

$$P(5 \text{ tails}) = \frac{C_{10,5}}{2^{10}} = \frac{252}{1024} \approx 0.246 \approx 24.6\%.$$

We will now try to explore the structure of the sample space of 10 coin flips. Based on the above, we know that there are 1024 members of this space. These will range from $HHHHHHHHHH$ to $TTTTTTTTTT$ and everything in between. A natural way to navigate this large space is to group outcomes together based on how many tails they contain.[1] A helpful way to understand the sample space is to group outcomes by how many tails they contain. Based on the 10^{th} row of Pascal's triangle, the sample space of 1024 outcomes can be organised as follows.

# of tails r	0	1	2	3	4	5	6	7	8	9	10
# of outcomes	1	10	45	120	210	252	210	120	45	10	1

The table above can be visualised as a bar chart:



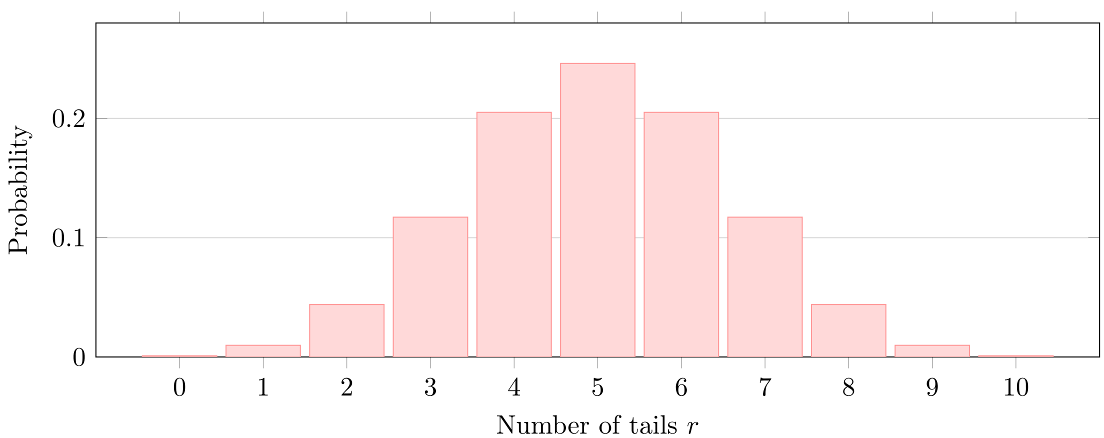
Here the x-axis consists of a discrete count: simply, it is the number of possible tails. The y-axis is the number of corresponding outcomes. So, in a sense, this graph is simply a visualisation of the 10^{th} row of Pascal's triangle.

[1]Of course, this could equally be done by starting with heads. But, since we have been using tails this whole time, we will commit to it.

Based on our discussion above, we know that the probabilities in this system can be calculated by dividing $C_{10,r}$ by $2^{10} = 1024$. Therefore, we can perform the division and then plot the results to see how *probability itself* is distributed amongst the outcomes in this system.

r **tails**	0	1	2	3	4	5	6	7	8	9	10
Probability (fraction)	$\frac{1}{1024}$	$\frac{10}{1024}$	$\frac{45}{1024}$	$\frac{120}{1024}$	$\frac{210}{1024}$	$\frac{252}{1024}$	$\frac{210}{1024}$	$\frac{120}{1024}$	$\frac{45}{1024}$	$\frac{10}{1024}$	$\frac{1}{1024}$
Approx. percentage	0.1%	1.0%	4.4%	11.7%	20.5%	24.6%	20.5%	11.7%	4.4%	1.0%	0.1%

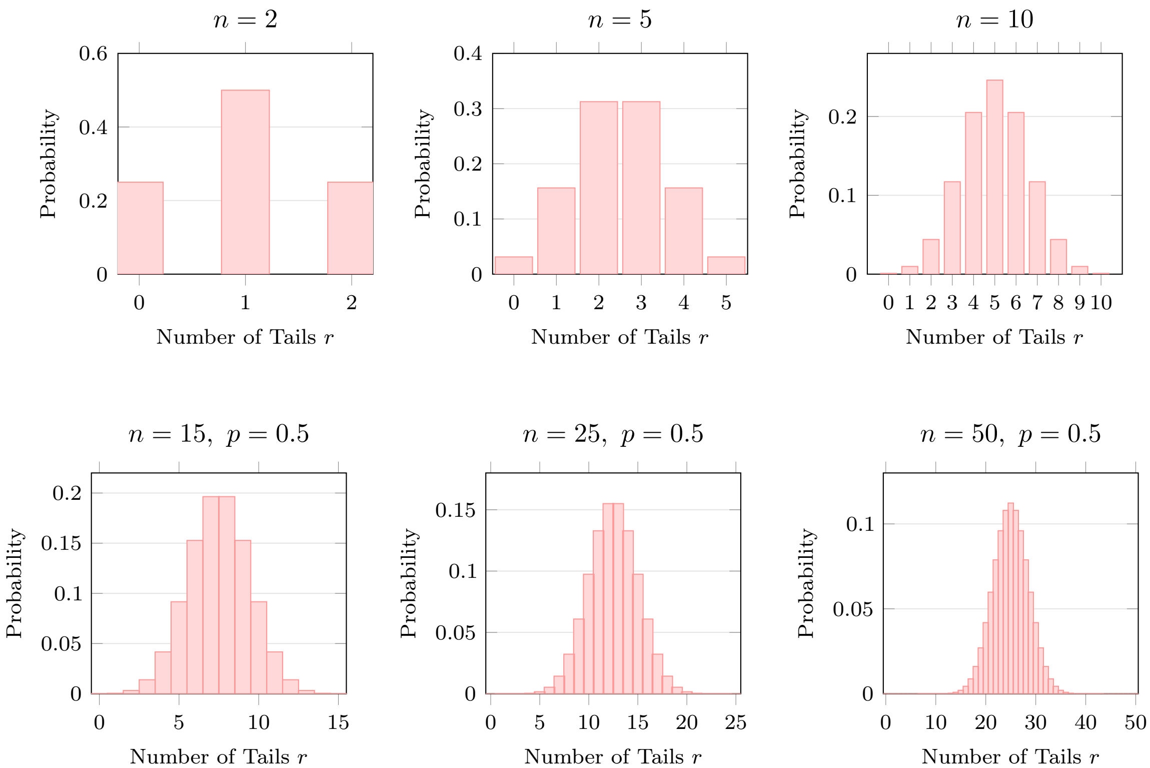
Observe the value of 24.6% that we obtain earlier. Using the above table, we can plot out this distribution of probabilities:



Notice that the shape is *precisely* the same as that of the bar chart from before. The only thing that has changed is the scale of the y-axis. This particular graph has a name: it is called a **binomial distribution**.

2.2 The binomial distribution for other values of n

Generally, we can create binomial distributions for other values of n, the total number of coin flips. The shape of these graphs will map out the way in which probability is distributed across the many, many possible outcomes of coin flips. The shape comes from the n^{th} row of Pascal's triangle, and after dividing by 2^n those numbers become probabilities. Generally speaking, if we were to keep making coin flips (i.e. so that n grows), then there will be more possible values of r, and the distribution becomes smoother and more concentrated around the middle (near $r = n/2$). We will speak about this "smoothness" in great detail in Lecture 10.



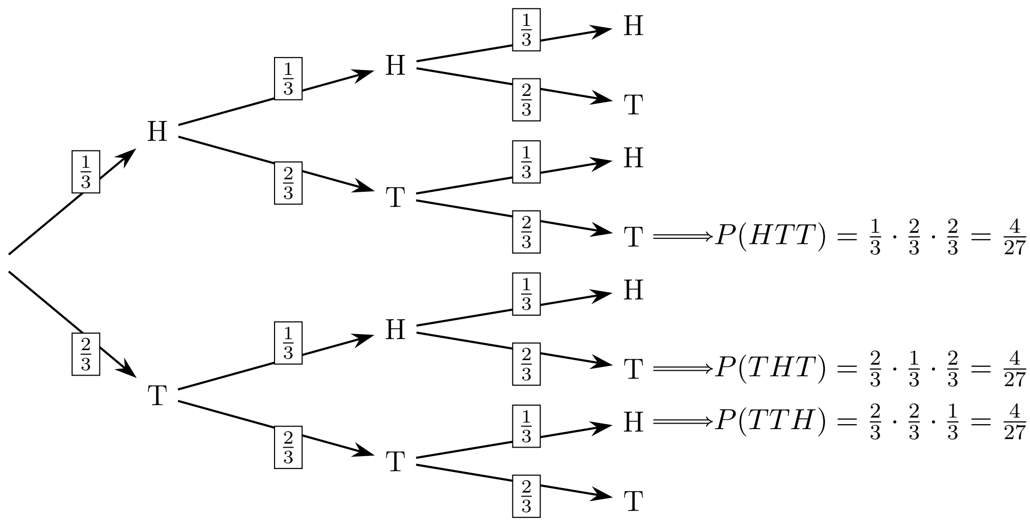
2.3 Example 2: flipping an unfair coin

Suppose now that we have a coin that has been tampered with in some way, so that the probabilities of landing on heads or tails are *no longer equally likely.* Let's assume that our "unfair" coin has probabilities:

$$P(H) = \frac{1}{3}, \qquad P(T) = \frac{2}{3}.$$

Question: what is the probability that we get exactly 2 tails in 3 flips?

In the fair coin case discussed above, all of the 8 outcomes are equally likely, so we could simply count up all of the favourable outcomes (there were 3) then divide by 8. However, for our unfair coin, it is *not* the case that all outcomes are equally likely. So, we cannot just count favourable outcomes and divide by 8. Instead, we can use the decision tree approach. The space of possible outcomes is the same as before, however, now the probabilities on each branch of the tree have changed:



Exactly 2 tails corresponds to the outcomes HTT, THT, TTH. These are mutually exclusive events, so our desired probability can be obtained by summing:

$$P(\text{exactly 2 tails}) = P(HTT \text{ or } THT \text{ or } TTH) = \frac{4}{27} + \frac{4}{27} + \frac{4}{27} = \frac{12}{27}.$$

Notice that the probability calculation of each of these outcomes is effectively the same: to calculate $P(HTT)$ or $P(THT)$ or $P(TTH)$, we simply multiply two copies of $P(T) = \frac{2}{3}$ together with one copy of $P(H) = \frac{1}{3}$. In fact, this style of calculation will work to calculate any further probabilities. For n-many flips, we have:

$$P(r \text{ many } T\text{'s}) = C_{n,r} \left(\frac{2}{3}\right)^r \left(\frac{1}{3}\right)^{n-r}.$$

Even more generally, if we were to adjust the probabilities of landing on tails or heads to other values, we could calculate the probability of r-many tails in n-many flips using the formula:

$$P(r \text{ many } T\text{'s in } n \text{ flips}) = C_{n,r}\,(P(T))^r\,(P(H))^{n-r}.$$

Here $(P(T))^r\,(P(H))^{n-r}$ is the probability of one specific ordering of r-many T's and $(n-r)$-many H's, and $C_{n,r}$ counts the number of different orderings of r-many T's and $(n-r)$-many H's.

3 The general theory

3.1 Definition of a binomial experiment

In Section 2 we saw two different versions of coin flips: one fair, and one "unfair" in the sense that the two possible outcomes had different probabilities. As a matter of fact, these are two examples of a more general process called a *binomial experiment*. There are four important defining features of a binomial experiment.

Definition: Binomial Experiment

A binomial experiment has the following properties:

(B1) There is a fixed number of trials. Call this number n.

(B2) The n trials are independent and repeated under identical conditions.

(B3) Each trial has only two outcomes: success (S) and failure (F).

(B4) On each trial the probability of success is the same. Call it p, and call the probability of failure q. Since each trial is either success or failure, $p + q = 1$, so $q = 1 - p$.

Put differently, a binomial experiment is a fixed number of independent trials, each with two possible outcomes and the same success probability. In a binomial experiment, the central problem is to determine the probability of r successes out of n trials. The assumption (B2) is used to simplify the calculation of these successive probabilities.

3.1.1 Example: fair coin flips

In our first example from the previous section:

(B1) n was the number of coin flips.

(B2) The coin flips are independent. In fact, they are not related to the outcomes of previous experiments at all (even physically).

(B3) Success was defined as "we got a tail" and failure was defined as "we did not get a tail".

(B4) For a fair coin, $p = P(T) = 0.5$ and $q = 0.5$.

For the central question of determining probabilities associated with this experiment, this is precisely what our previous discussion was doing. We already saw that

$$P(r \text{ tails out of } n \text{ flips}) = \frac{C_{n,r}}{2^n}$$

for a fair coin.

For the unfair coin, the binomial setup is essentially the same, except that we adjusted the probabilities to $p = \frac{2}{3}$ and $q = \frac{1}{3}$. We saw previously that the associated probabilities for different values of r can be calculated with the formula:

$$P(r \text{ many } T\text{'s}) = C_{n,r} \left(\frac{2}{3}\right)^r \left(\frac{1}{3}\right)^{n-r} = C_{n,r} \cdot P(T)^r \cdot P(H)^{n-r}.$$

General Formula for Binomial Probability

For a binomial experiment with parameters n, p, $q = 1 - p$, the probability of getting exactly r successes is

$$P(r \text{ successes out of } n \text{ trials}) = C_{n,r}\, p^r\, q^{n-r}.$$

Here:

- p^r is the success part of the probability for one specific order.
- q^{n-r} is the failure part of the probability for that same specific order.
- $p^r q^{n-r}$ is the probability of one particular sequence that has r successes and $(n - r)$ failures.
- $C_{n,r}$ counts how many different sequences have r successes and $(n - r)$ failures.

3.2 What does p do to the distribution?

Generally speaking, we can use the formula

$$P(r \text{ successes out of } n \text{ trials}) = C_{n,r}\, p^r\, q^{n-r}.$$

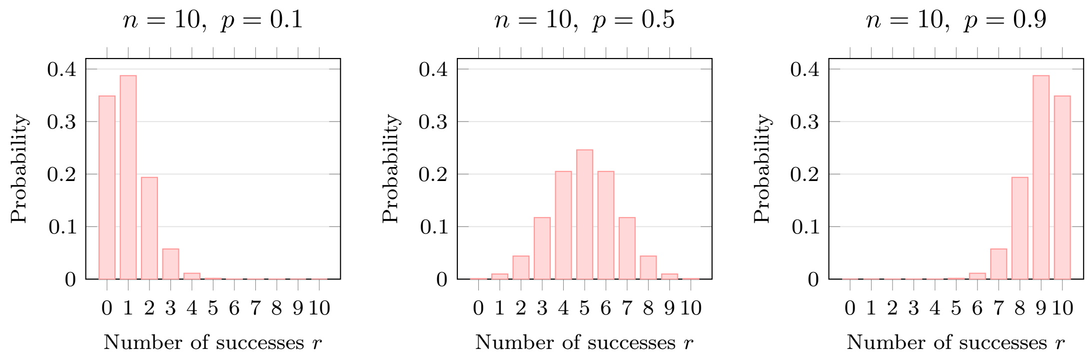
to describe a binomial distribution associated to every binomial experiment. Simply: we take the x-axis to be the possible values of the random variable (from 0 to n), and we take the y-axis to be the output of the formula above.

In the case of a fair coin, the success probability $p = 0.5$, which is exactly equal to q. This causes the resulting binomial distribution to be perfectly symmetric. Notice this in the graphs detailed in Section 2. If we start to adjust the success probability away from $p = 0.5$, then this will cause the distribution to lose its symmetry.

For illustration purposes, let's suppose that we are working with a binomial experiment of $n = 10$ trials. In the case that the probability $p = 0.5$, the resulting distribution is the figure in Section 2.1. We will now briefly discuss what happens if we make p bigger or smaller.

- If p is near 0, say $p = 0.1$, the chance of getting a successful outcome to any trial is much lower. Therefore, the chance of getting a large number of successes, such as $7, 8, 9$ or 10 successful outcomes in 10 trials is *extremely low*. Most likely, you will only get 0, 1 or 2 successful outcomes in 10 attempts. As such, the binomial distribution associated to a system like this will actually have the majority of its probability distributed towards the left side of the x-axis, closer to the values of $0, 1$ and 2.
- If p is near 1, say $p = 0.9$, then it would now be *extremely likely* to receive a successful trial. Therefore, it would be very unlucky to receive only one or two successful outcomes in 10 attempts, and instead you would expect to get somewhere around $7, 8, 9$ or 10 successes. On the level of the binomial distribution, having such a high chance of success will push the distribution to the *right* and cause most of the probability to be centred around the $8, 9, 10$ area of the x-axis.

The graphs below depict the actual binomial distributions for $n = 10$ trials with success probabilities of 0.1, 0.5 and 0.9, respectively.



3.3 Expected value and variance for the binomial distribution

For a binomial distribution, there are two especially important formulas:

Binomial expected value and variance

If X counts the number of successes in a binomial experiment with parameters n, p and $q = 1-p$, then

$$E(X) = np, \qquad \mathrm{Var}(X) = npq.$$

These formulas are useful because they let us quickly understand the centre and spread of a binomial distribution without calculating the entire distribution by hand.

Technically, we could verify the above two formulas using some algebra. However, we will not do this and will instead argue loosely: imagine the n trials as n slots:

$$\square \quad \square \quad \square \quad \cdots \quad \square$$

Each slot contributes either 1 success or 0 successes. Since the probability of success in each slot is p, according to our Frequentist interpretation of probability, each slot contributes a long-term expected average of p successes. Since there are n slots, the total expected number of successes is

$$p + p + \cdots + p = np,$$

which recovers the formula $E(X) = np$. As for the variance formula, we will not derive this here. However, we will note that the formula

$$\mathrm{Var}(X) = npq$$

tells us how volatile the number of successes is:

- If p is close to 0 or close to 1, then the experiment is less volatile and more tightly packed around $E(X)$, because the outcome is more predictable.
- If p is near 0.5, then the experiment is more volatile, because success and failure are both much more plausible on each trial.

4 Exercises

4.1 The roulette wheel

Suppose that there are 20 students in our class. You study hard and after a few years you successfully graduate from LUJ. Dr. O'Connell is so proud of you that he decides to take you all to the casino in order to play on the roulette wheel. Each student spins the wheel once.

There are 36 numbered slots on the roulette wheel, and the green slot makes 37 total outcomes. Dr. O'Connell gives you the following offer:

> You get \$50,000 if the ball lands on the green slot, and \$0 if any other outcome happens.

The major question we will answer is:

Exercise

What is the probability that exactly 3 students will win?

To solve this question, we will use the general binomial formula mentioned in Section 3.2. However, before we start to calculate anything, we must first confirm that this situation is actually a binomial experiment to begin with.

- B1: There are 20 students and they each try the roulette wheel once, so $n = 20$ trials.
- B2: Assuming that each spin is independent and the roulette wheel is fair, the trials are indeed independent from each other.
- B3: There are only two possible outcomes to each trial:

$$\text{Success} = \text{lands on green}, \qquad \text{Failure} = \text{lands elsewhere}.$$

- B4: each slot of the roulette wheel has an equally likely chance of being the outcome. Since there are 37 slots on the wheel and only one green slot to land on, the probabilities of success and failure are:

$$p = P(\text{success}) = \frac{1}{37}, \qquad q = P(\text{failure}) = 1 - p = \frac{36}{37}.$$

Now that we have confirmed our system can be seen as a binomial experiment, we can solve the main question.

Solution

We use the binomial formula:

$$P(r \text{ successes out of } n \text{ trials}) = C_{n,r}\, p^r\, q^{n-r}.$$

In this case, $n = 20$ since there are 20 students, $r = 3$, $p = \frac{1}{37}$ and $q = \frac{36}{37}$. We substitute all of these values into the formula and see:

$$P(\text{exactly 3 wins}) = C_{20,3}\left(\frac{1}{37}\right)^3\left(\frac{36}{37}\right)^{17} = \frac{20!}{3!\cdot 17!}\left(\frac{1}{37}\right)^3\left(\frac{36}{37}\right)^{17}.$$

Numerically, this is approximately

$$P(\text{exactly 3 wins}) \approx 0.014 \approx 1.4\%.$$

As an alternate explanation of this probability: we are performing 20 independent trials and we want 3 successes and 17 failures. One particular outcome that represents this could be:

$$S,\ S,\ S,\ F,\ F,\ \ldots,\ F,$$

i.e. the first 3 students are successful and then the 17 students after that all fail. According to the assumption (B2), all of these trials are independent from each other. So, the probability associated with this outcome can be calculated by simply multiplying the probability of each outcome. We have:

$$P(S \text{ then } S \text{ then } S \text{ then } F \text{ then } \ldots \text{ then } F) = \frac{1}{37}\cdot\frac{1}{37}\cdot\frac{1}{37}\cdot\frac{36}{37}\cdot\ldots\cdot\frac{36}{37},$$

which is the same as writing p^3q^{17}. This is the formula to calculate a single outcome that corresponds to 3 successes and 17 failures. But, there are $C_{20,3}$-many different ways to place the 3 successes among 20 trials. These sequences correspond to mutually exclusive outcomes, so we add their probabilities: producing $C_{20,3}p^3q^{17}$.

$$P(\text{exactly 3 wins}) = \underbrace{p^3q^{17} + p^3q^{17} + \ldots + p^3q^{17}}_{C_{20,3}\text{-many times}} = C_{20,3}\cdot p^3q^{17},$$

which is the same formula as before.

4.2 Loot boxes

A game has a rare loot item with a 0.26% chance of appearing on each play. Each play costs \$2.49. We will treat "success" as getting the loot.

Questions

(a) What is the probability of getting exactly 2 wins in 100 attempts?
(b) What is the probability of getting exactly 3 wins in 1000 attempts?
(c) How much money should you expect to pay before you find the loot?

Solution

The probability of success is

$$p = 0.0026, \qquad q = 1 - p = 0.9974.$$

(a) Here $n = 100$ and $r = 2$. So:

$$P(X = 2) = C_{100,2}(0.0026)^2(0.9974)^{98}.$$

Numerically, this is approximately

$$P(X = 2) \approx 0.0259 = 2.59\%.$$

(b) Here $n = 1000$ and $r = 3$. So:

$$P(X = 3) = C_{1000,3}(0.0026)^3(0.9974)^{997}.$$

Numerically, this is approximately

$$P(X = 3) \approx 0.218 = 21.8\%.$$

(c) We are looking for the number of attempts n so that the expected number of wins reaches 1. Since $E(X) = np$, we solve:

$$1 = n(0.0026) \qquad \Longrightarrow \qquad n = \frac{1}{0.0026} \approx 384.6.$$

So, on average, we would expect to need about 385 plays for one loot item. Since each play costs \$2.49, this would cost

$$385 \times 2.49 = \$958.65.$$

4.3 Kaitenzushi

A kaitenzushi restaurant allows you to play a game and win a prize after you insert 5 plates into the machine. The probability of winning the prize is 25%.

Questions

(a) What is the probability of playing 5 times and winning twice?
(b) What is the probability of winning 2 times in 2 attempts?
(c) On average, how many plates should you expect to eat before winning a prize?

Solution

Success is defined as winning the prize. Then

$$p = 0.25, \qquad q = 0.75.$$

(a) Here $n = 5$ and $r = 2$, so:

$$P(X = 2) = C_{5,2}(0.25)^2(0.75)^3.$$

Numerically,

$$P(X = 2) = 10(0.0625)(0.421875) = 0.263671875 \approx 26.4\%.$$

(b) Winning 2 times in 2 attempts means winning both attempts. Since the attempts are independent:

$$P(\text{win then win}) = 0.25 \times 0.25 = 0.0625 = 6.25\%.$$

(c)

(d) Since $E(X) = np$, we want $E(X) = 1$, so:

$$1 = n(0.25) \qquad \Longrightarrow \qquad n = 4.$$

So we expect about 4 attempts before winning one prize. Since each attempt requires 5 plates, this corresponds to

$$4 \times 5 = 20$$

plates on average.

4.4 Rolling a dice

Imagine you roll a standard 6-sided die three times.

We will treat "success" as rolling a 6. Then each roll has only two outcomes (6 or not 6), the rolls are independent, and the success probability is constant:

$$n = 3, \qquad p = P(\text{roll a } 6) = \frac{1}{6}, \qquad q = \frac{5}{6}.$$

Question

What is the probability of getting exactly two 6's?

Solution

To get exactly two 6's, we need $r = 2$ successes out of $n = 3$ trials. The binomial formula gives

$$P(r \text{ successes}) = C_{n,r} p^r q^{n-r}.$$

So

$$P(\text{exactly two 6's}) = C_{3,2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right) = 3 \cdot \frac{1}{36} \cdot \frac{5}{6} = \frac{15}{216} = \frac{5}{72} \approx 6.9\%.$$

Narrative interpretation: there are $C_{3,2} = 3$ ways to choose which two of the three rolls are the 6's, and each such pattern has probability $\left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)$.

4.5 The basketball cat

A cat is shooting basketball free throws. Cats are not very good at basketball, so suppose the probability a shot goes in is $p = 0.5$. The cat takes 6 shots.

We treat "success" as "the shot goes in". Then:

$$n = 6, \qquad p = 0.5, \qquad q = 0.5.$$

Questions

(a) What is the probability that the cat makes exactly 4 shots?
(b) What is the probability that the cat makes at least 4 shots?

Solution

Let X be the number of made shots.

(a) Exactly 4 makes means $r = 4$ successes out of $n = 6$ trials:

$$P(X = 4) = C_{6,4}(0.5)^4(0.5)^2 = C_{6,4}(0.5)^6 = \frac{C_{6,4}}{64} = \frac{15}{64} \approx 23.4\%.$$

(b) "At least 4" means 4 or 5 or 6 makes, so we add those mutually exclusive cases:

$$\begin{aligned} P(X \geq 4) &= P(X = 4) + P(X = 5) + P(X = 6) \\ &= \frac{C_{6,4}}{64} + \frac{C_{6,5}}{64} + \frac{C_{6,6}}{64} = \frac{15 + 6 + 1}{64} = \frac{22}{64} = \frac{11}{32} \approx 34.4\%. \end{aligned}$$