Many Coin Flips

In this section we are going to consider the number of heads on n coin flips. This thought experiment is going to be a basis for much probability theory! It goes far beyond coin flips.

Say a coin comes up heads with probability p. Most coins are fair and as such come up heads with probability $p = 0.5$. There are many events for which coin flips are a great analogy that have different values of p so let's leave p as a variable. You can try simulating flipping coins here. Note that **H** is short for Heads and **T** is short for Tails. We think of each coin as distinct:

Coin Flip Simulator

Number of flips n: `10`   Probability of heads p: `0.60`   New simulation

Simulator results:

> H, T, H, T, H, H, T, H, T, H

Total number of heads: 6

Using the math in this chapter we will be able to show that the probability of getting exactly 6 heads (when the probability of heads on each flip is 0.6) = 0.251. Specifically in this chapter:

1. Warmups: We calculate the probability of a few exact outcomes.
2. Exactly k heads. We derive the general formula.
3. More than k heads. We explore this interesting related problem.

Warmups

In all of these warmups we are going to consider the probability of different outcomes when you flip a coin n times and each time the probability of heads is p. In each solution we will consider the case where $n = 10$ and $p = 0.6$.

What is the probability that all n flips are heads?

This question is asking what is the probability of getting the outcome:

> H, H, H, H, H, H, H, H, H, H

Where each flip lands in heads (**H**). Each coin flip is independent so we can use the rule for probability of and with independent events. As such, the probability of n heads is p multiplied by itself n times: p^n.

If $n = 10$ and $p = 0.6$ then the probability of n heads = $p^n = 0.6^{10} \approx 0.006$.

What is the probability that all n flips are tails?

Lets say $n = 10$ this question is asking what is the probability of getting:

> T, T, T, T, T, T, T, T, T, T

Each coin flip is independent. The probability of tails on any coin flip is $1 - p$. Again, since the coin flips are independent, the probability of tails n times on n flips is $(1 - p)$ multiplied by itself n times: $(1 - p)^n$. If $n = 10$ and $p = 0.6$ then the probability of n tails is around 0.0001.

First k heads then $n - k$ tails

> Lets say $n = 10$ and $k = 4$, this question is asking what is the probability of getting:
>
> H, H, H, H, T, T, T, T, T, T
>
> The coins are still independent! The first k heads occur with probability p^k the run of $n - k$ tails occurs with probability $(1 - p)^{n-k}$. The probability of k heads then $n - k$ tails is the product of those two terms: $p^k \cdot (1 - p)^{n-k}$

Exactly k heads

Next lets try to figure out the probability of exactly k heads in the n flips. Importantly we don't care where in the n flips that we get the heads, as long as there are k of them. Note that this question is different than the question of first k heads and then $n - k$ tails which requires that the k heads come first! That particular result does generate exactly k coin flips, but there are others.

There are many others! Lets ask the computer to list the ways we could generate exactly $k = 4$ heads within $n = 10$ coin flips. The output region is scrollable:

```
(H, H, H, H, T, T, T, T, T, T)
(H, H, H, T, H, T, T, T, T, T)
(H, H, H, T, T, H, T, T, T, T)
(H, H, H, T, T, T, H, T, T, T)
(H, H, H, T, T, T, T, H, T, T)
(H, H, H, T, T, T, T, T, H, T)
(H, H, H, T, T, T, T, T, T, H)
(H, H, T, H, H, T, T, T, T, T)
(H, H, T, H, T, H, T, T, T, T)
(H, H, T, H, T, T, H, T, T, T)
(H, H, T, H, T, T, T, H, T, T)
(H, H, T, H, T, T, T, T, H, T)
(H, H, T, H, T, T, T, T, T, H)
(H, H, T, T, H, H, T, T, T, T)
(H, H, T, T, H, T, H, T, T, T)
(H, H, T, T, H, T, T, H, T, T)
(H, H, T, T, H, T, T, T, H, T)
(H, H, T, T, H, T, T, T, T, H)
```

Let's call each of these rows an "ordering" since each row is a unique way to order the 4 heads and 6 tails. Let N be the number of unique orderings.

Exactly how many unique orderings are there with $k = 4$ heads in $n = 10$ flips? $N = 210$. Why? Each ordering listed above is a permutation of the list [H, H, H, H, T, T, T, T, T, T]. As such, the question "how many orderings are there?" is analogous to the question, how many distinct orderings of characters are possible for the string "HHHHTTTTTT"? Because H characters are indistinct from one another (and same for T) we can solve this problem using [Permutations of Indistinct Objects](#):

$$\text{Num Orderings} = \frac{n!}{k!(n-k)!} = \binom{n}{k}$$

Each of these orderings can be thought of as one event, or outcome, of flipping 10 coins. Let's name E_i to be the event that we get the exact outcome in the ith row. E_1 is the event that we get [H, H, H, H, T, T, T, T, T, T]. E_2 is the event that we get [H, H, H, T, H, T, T, T, T, T] and so on.

The probability of exactly $k = 4$ heads is the probability of the **or** of each of these events E_i. If you flip a coin 10 times, it is not possible to have more than one of the E_i events occur (eg both E_1 and E_2 can't be true if you flip 10 coins). In other words, each of these events E_i is "mutually exclusive" and as such:

$$P(\text{exactly } k \text{ heads}) = P(E_1 \text{ or } E_2 \text{ or } \ldots \text{ or } E_N)$$
$$= \sum_{i=1}^{N} P(E_i)$$

The next question is, what is the probability of each of these events E_i?

Here is a arbitrarily chosen ordering which satisfies the event of exactly $k = 4$ heads in $n = 10$ coin flips. It is E_{128}, the ordering on row 128 in the list above:

T, H, T, T, H, T, T, H, H, T

What is the probability of event E_{128}, the exact sequence of heads and tails in the example above? Each coin flip is still independent, so we multiply p for each heads and $1 - p$ for each tails.

$$P(E_{128}) = (1 - p) \cdot p \cdot (1 - p) \cdot (1 - p) \cdot p \cdot (1 - p) \cdot (1 - p) \cdot p \cdot p \cdot (1 - p)$$

If you rearrange these multiplication terms you get:

$$P(E_{128}) = p \cdot p \cdot p \cdot p \cdot (1 - p) \cdot (1 - p) \cdot (1 - p) \cdot (1 - p) \cdot (1 - p) \cdot (1 - p)$$
$$= p^4 \cdot (1 - p)^6$$

There is nothing too special about row 128. If you chose any row, you would get k independent heads and $n - k$ independent tails. For any row i, $P(E_i) = p^k \cdot (1 - p)^{n-k}$.

Now we are ready to calculate the probability of exactly k heads:

$$P(\text{exactly } k \text{ heads}) = \sum_{i=1}^{N} P(E_i) \qquad \text{Mutual Exclusion}$$
$$= \sum_{i=1}^{N} p^k \cdot (1 - p)^{n-k} \qquad \text{Sub in } P(E_i)$$
$$= N \cdot p^k \cdot (1 - p)^{n-k} \qquad \text{Sum } N \text{ times}$$
$$= \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \qquad \text{Perm of indistinct objects}$$

Let's bring that back to the example of getting exactly $k = 4$ heads out of $n = 10$ where the probability of getting a heads on any one flip is $p = 0.6$. The probability of exactly k heads is:

$$P(\text{exactly } k \text{ heads}) = \binom{n}{k} \cdot p^k \cdot (1 - p)^{n-k} \qquad \text{Just derived!}$$
$$= \binom{10}{4} \cdot 0.6^4 \cdot (1 - 0.6)^{10-4} \qquad \text{Sub in } n, k, p$$
$$= 210 \cdot 0.6^4 \cdot 0.4^6 \qquad \text{Simplify}$$
$$= 0.111$$

We are done! This result is truly useful. Care about the probability of number of voters who vote for a candidate? Care about the number of people who get a disease? Care about the number of people who click on an ad? This derivation is the basis for all of that! Later in this book we will formalize this incredible result into the Binomial Random Variable.

More than k heads

There are many cases where you care about the probability of more than k heads. We can derive a solution based on the forumla for exactly k heads.

If you want the probability of getting more than $k = 4$ heads out of $n = 10$ coin flips it is the probability of the **or** of each of the events where you get exactly i heads for $i = 5, 6, 7, 8, 9, 10$. Note that all the events where you get exactly i heads are mutually exclusive from the events where you get exactly j heads when $i \neq j$. For example if you flip a coin 10 times, it is not possible to get exactly 4 heads and to also get exactly 5 heads. As such the probability of **or** becomes addition:

$$P(\text{more than } k \text{ heads}) = \sum_{i=k+1}^{n} P(\text{exactly } i \text{ heads}) \qquad \text{Mutual Exclusion}$$

$$= \sum_{i=k+1}^{n} \binom{n}{i} \cdot p^i \cdot (1-p)^{n-i} \qquad \text{Substitution}$$