EXPECTED VALUE

DISCRETE STRUCTURES II

DARRYL HILL

BASED ON THE TEXTBOOK:

DISCRETE STRUCTURES FOR COMPUTER SCIENCE: COUNTING, RECURSION, AND PROBABILITY

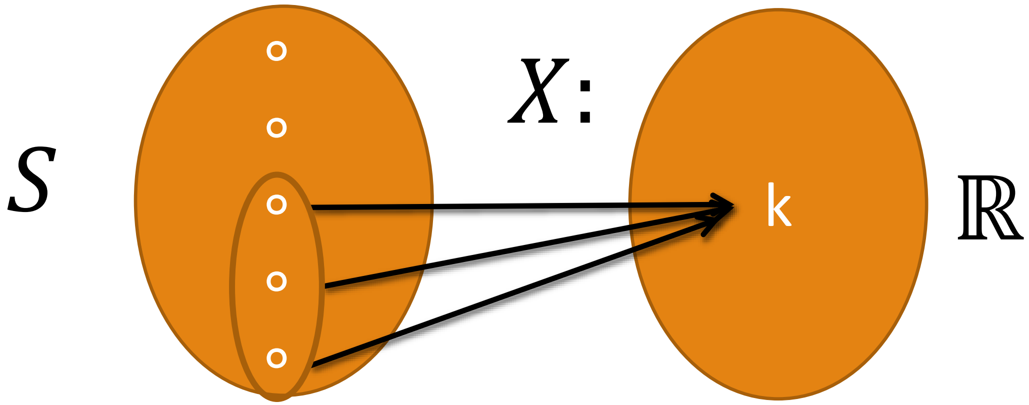
BY MICHIEL SMID

S: Sample space
Outcome: Element of S
Event: Subset of S
$\Pr(x) : x \in S \rightarrow [0,1]$
$\sum_{w \in S} \Pr(w) = 1$

Random Variable:

function $X: S \rightarrow \mathbb{R}$

"neither random nor variable"



$$S = \{a, b, c\}$$
$$\Pr(a) = \frac{4}{5}, \Pr(b) = \frac{1}{10}, \Pr(c) = \frac{1}{10}$$

$$X(a) = 1, X(b) = 2, X(c) = 3$$

$E(X) =$ expected value of X

"Average" value of X would be?

First instinct $E(X) = \frac{1+2+3}{3} = 2$

But we choose $a \in S$ 80 % of the time, and the other two outcomes 10% of the time each! What is the average then if we consider the probabilities?

S: Sample space

Outcome: Element of S

Event: Subset of S

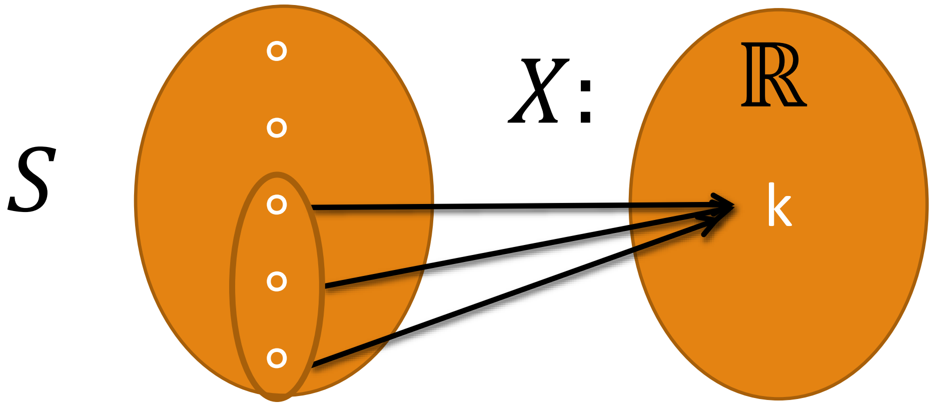
$\Pr(x) : x \in S \rightarrow [0,1]$

$\sum_{w \in S} \Pr(w) = 1$

Random Variable:

function $X: S \rightarrow \mathbb{R}$

"neither random nor variable"



$$S = \{a, b, c\}$$

$$\Pr(a) = \frac{4}{5}, \Pr(b) = \frac{1}{10}, \Pr(c) = \frac{1}{10}$$

$$X(a) = 1, X(b) = 2, X(c) = 3$$

$E(X)$ = expected value of X - "Weighted Average“
Where each value of a random variable is given a weight proportional to its probability.

$$E(X) = 1 \cdot \frac{4}{5} + 2 \cdot \frac{1}{10} + 3 \cdot \frac{1}{10} = \frac{13}{10}$$

Note that:

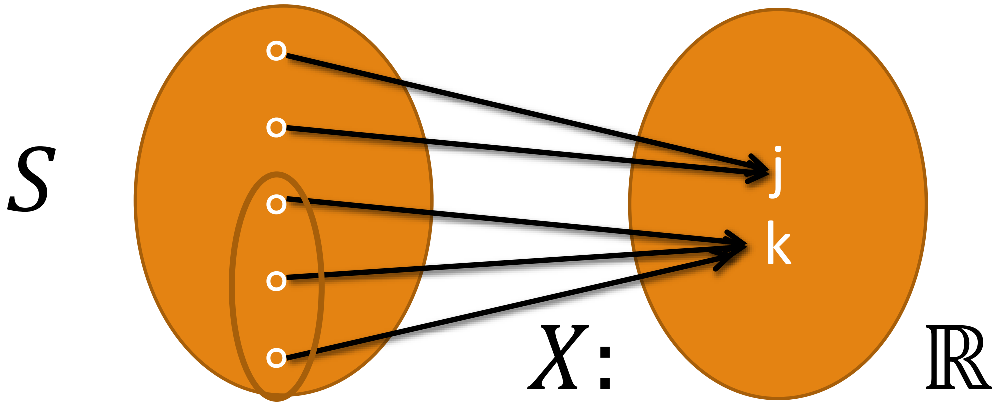
$$\sum weights = \sum_{w \in S} \Pr(w) = 1$$

S: Sample space
Outcome: Element of S
Event: Subset of S
$\Pr(x) : x \in S \rightarrow [0,1]$
$\sum_{w \in S} \Pr(w) = 1$

Random Variable:

function $X: S \rightarrow \mathbb{R}$

"neither random nor variable"



$$S = \{a, b, c\}$$

$$\Pr(a) = \frac{4}{5}, \Pr(b) = \frac{1}{10}, \Pr(c) = \frac{1}{10}$$

$$X(a) = 1, X(b) = 2, X(c) = 3$$

The definition of expected value is:

$$E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)$$

For the random variable X above:

$$E(X) = X(a) \cdot \Pr(a) + X(b) \cdot \Pr(b) + X(c) \cdot \Pr(c)$$

$$E(X) = 1 \cdot \frac{4}{5} + 2 \cdot \frac{1}{10} + 3 \cdot \frac{1}{10} = \frac{13}{10}$$

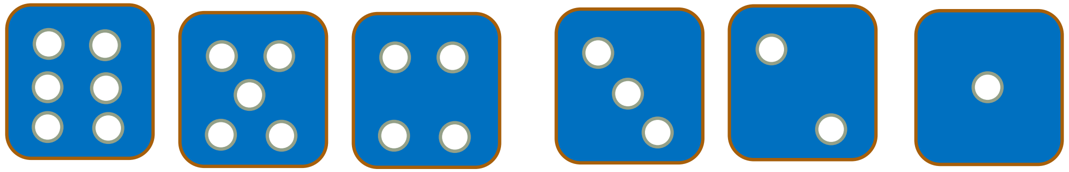
$$\boxed{E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)}$$

Roll fair die: $S = \{1,2,3,4,5,6\}$, uniform probability.

Thus every element of S has probability $\frac{1}{6}$.

X = result of roll; $X(i) = i$

$$E(X) = \sum_{i=1}^{6} X(i) \cdot \Pr(i)$$



$$E(X) = 1 \cdot \Pr(1) + 2 \cdot \Pr(2) + 3 \cdot \Pr(3) + 4 \cdot \Pr(4) + 5 \cdot \Pr(5) + 6 \cdot \Pr(6)$$

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

$$= 3.5$$

In this particular case it is simply the average die roll – not the value you would expect to see (since you cannot roll 3.5).

You can think of it as rolling a die many times (say millions of times), and taking the average of all rolls. The average would be close to 3.5.

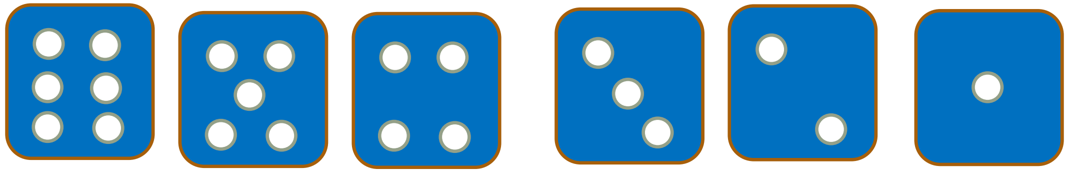
$$E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)$$

Roll fair die: $S = \{1,2,3,4,5,6\}$, uniform probability

X = result of roll; $X(i) = i$

$$E(X) = 3.5$$

$$Y = \frac{1}{result}$$



$$Y(i) = \frac{1}{i}$$

$$E(Y) = \sum_{i=1}^{6} \frac{1}{i} \cdot \frac{1}{6}$$

$$= \frac{1}{6} \cdot \sum_{i=1}^{6} \frac{1}{i}$$

$$E(Y) = \frac{1}{6} \cdot \left(\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6}\right)$$

$$= \frac{1}{6} \cdot \left(\frac{120}{120} + \frac{60}{120} + \frac{40}{120} + \frac{30}{120} + \frac{24}{120} + \frac{20}{120}\right)$$

$$= \frac{49}{120}$$

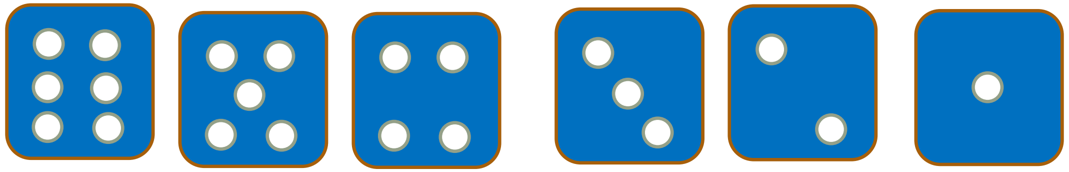
$$\boxed{E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)}$$

Roll fair die: $S = \{1,2,3,4,5,6\}$, uniform probability

X = result of roll; $X(i) = i$

$$E(X) = 3.5$$

$$Y = \frac{1}{result}$$



$$Y(i) = \frac{1}{i}$$

$$Y = \frac{1}{X}$$

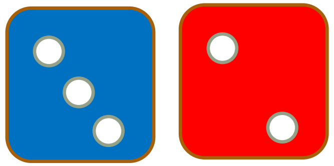
$$E(Y) = \frac{49}{120} = 0.408\overline{3} = E\left(\frac{1}{X}\right)$$

$$E(X) = \frac{7}{2},$$

$$\frac{1}{E(X)} = \frac{2}{7} \approx 0.286$$

Thus in general $E\left(\frac{1}{X}\right) \neq \frac{1}{E(X)}$

$$E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)$$



Roll fair red die and fair blue die.

$$S = \{\, (i,j) \mid 1 \le i \le 6, 1 \le j \le 6 \,\}$$

Uniform probability $\quad \Pr(i,j) = \dfrac{1}{36}$

$X: S \to \mathbb{R}$: X = red + blue: $X(i,j) = i + j$

We will look at 3 ways to compute the **Expected Value**. We will go in order of difficulty.

	1	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$\boxed{E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)}$$

Roll fair red die and fair blue die.

$$S = \{ (i,j) \mid 1 \le i \le 6, 1 \le j \le 6 \}$$

Uniform probability: $\Pr(i,j) = \frac{1}{36}$

$X: S \to \mathbb{R}$: X = red + blue: $X(i,j) = i + j$

$$E(X) = \sum_{(i,j) \in S} X(i,j) \cdot \Pr(i,j)$$

$$= \frac{1}{36} \sum_{(i,j) \in S} X(i,j)$$

$$\sum_{(i,j) \in S} X(i,j) = \text{sum of all entries}$$

$$= 252$$

$$E(X) = \sum_{(i,j) \in S} X(i,j) \cdot \Pr(i,j) = \frac{252}{36} = 7$$

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)$$

Roll fair red die and fair blue die.

$$S = \{ (i,j) \mid 1 \le i \le 6, 1 \le j \le 6 \}$$

Uniform probability: $\Pr(i,j) = \frac{1}{36}$

$X: S \to \mathbb{R}$: $X =$ red + blue: $X(i,j) = i + j$

Goal is to get a different formula that is shorter and easier.

If we look at the table (which is really just the function $X(i,j)$), there are entries that occur multiple times.

For instance, 4 occurs 3 times.

$$X = 4 \leftrightarrow \{(3,1),(2,2),(1,3)\}$$

We only look at elements of the summation where $X = 4$. There are 3, so the probability sums to

$$\frac{3}{36}$$

	1	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)$$

Roll fair red die and fair blue die.

$$S = \{ (i,j) \mid 1 \le i \le 6, 1 \le j \le 6 \}$$

Uniform probability: $\Pr(i,j) = \frac{1}{36}$

$X : S \to \mathbb{R}$: X = red + blue: $X(i,j) = i + j$

Of course this is the definition of an event, and this is exactly how we compute the probability of an event.

The event $X = 4 \leftrightarrow \{(3,1),(2,2),(1,3)\}$

$$\Pr(X = 4) = \frac{|\{(3,1),(2,2),(1,3)\}|}{|S|}$$

$$= \frac{3}{36}$$

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)$$

Roll fair red die and fair blue die.

$$S = \{ (i,j) \mid 1 \le i \le 6, 1 \le j \le 6 \}$$

Uniform probability: $\Pr(i,j) = \frac{1}{36}$

$X: S \to \mathbb{R}$: $X =$ red + blue: $X(i,j) = i + j$

We can rewrite our summation then to sum over all possible values of X and assign each of these **Events** the appropriate weight.

X can take on all values from 2 up to 12.

$E(X)$

$= 2 \cdot \quad + 3 \cdot \quad + 4 \cdot \quad + 5 \cdot \quad + 6 \cdot$

$+ 7 \cdot \quad + 8 \cdot \quad + 9 \cdot \quad + 10 \cdot \quad + 11 \cdot$

$+ 12 \cdot$

	1	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)$$

Roll fair red die and fair blue die.

$$S = \{\, (i,j) \mid 1 \le i \le 6, 1 \le j \le 6 \,\}$$

Uniform probability: $\Pr(i,j) = \frac{1}{36}$

$X: S \to \mathbb{R}$: $X =$ red + blue: $X(i,j) = i + j$

We can rewrite our summation then to sum over all possible values of X and assign each of these **Events** the appropriate weight.

X can take on all values from 2 up to 12.

$$E(X) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + 5 \cdot \frac{4}{36} + 6 \cdot \frac{5}{36} + 7 \cdot \frac{6}{36} + 8 \cdot \frac{5}{36} + 9 \cdot \frac{4}{36} + 10 \cdot \frac{3}{36} + 11 \cdot \frac{2}{36} + 12 \cdot \frac{1}{36} = 7$$

	1	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

$$E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)$$

Roll fair red die and fair blue die.

$$S = \{\, (i,j) \mid 1 \le i \le 6, 1 \le j \le 6 \,\}$$

$X: S \to \mathbb{R}$: $X =$ red + blue: $X(i,j) = i + j$

Still pretty painful. But using the former method, there are 36 entries to add up.

Using this method there are 11 entries to add up.

$$E(X) = 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + 5 \cdot \frac{4}{36} + 6 \cdot \frac{5}{36} + 7 \cdot \frac{6}{36} + 8 \cdot \frac{5}{36} + 9 \cdot \frac{4}{36} + 10 \cdot \frac{3}{36} + 11 \cdot \frac{2}{36} + 12 \cdot \frac{1}{36} = 7$$

	1	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

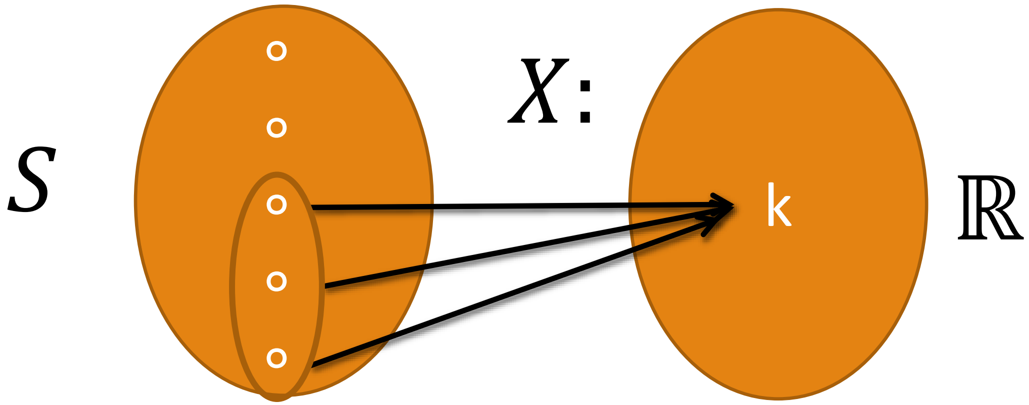
Random Variable: $X: S \to \mathbb{R}$

Event: "$X = k$"$\leftrightarrow \{w \in S: X(w) = k\}$

$k \in$ range of function X.

Instead of looking at every element of S, we look at every **Event** defined by the range of X, that is, all values X can take.

Sum over **Events** instead of **Outcomes**



$$E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)$$

Gather all elements $w \in S$ for which $X(w) = k$. Then the above summation can be rewritten as:

$$E(X) = \sum_{\forall\, k} \quad \sum_{w:X(w)=k} X(w) \cdot \Pr(w)$$
$$= \sum_{\forall\, k} \quad \sum_{w:X(w)=k} k \cdot \Pr(w)$$

We are still summing over all elements of S but we are dividing into subsets based on the Events $X = k$.

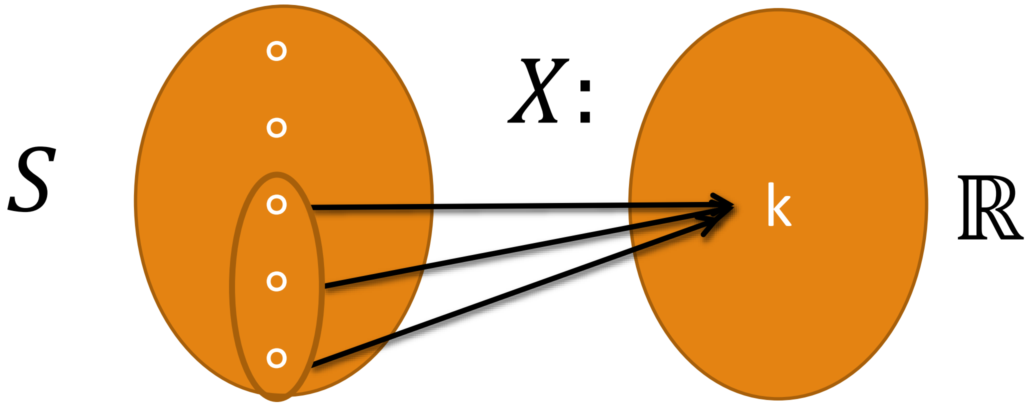
Random Variable: $X: S \to \mathbb{R}$

Event: "$X = k$"$\leftrightarrow \{w \in S: X(w) = k\}$

$k \in$ range of function X.

Instead of looking at every element of S, we look at every **Event** defined by the range of X, that is, all values X can take.

Sum over **Events** instead of **Outcomes**



$$E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)$$

Gather all elements $w \in S$ for which $X(w) = k$. Then the above summation can be rewritten as:

$$E(X) = \sum_{\forall\, k} \sum_{w: X(w)=k} X(w) \cdot \Pr(w)$$

$$= \sum_{\forall\, k} k \cdot \left(\sum_{w: X(w)=k} \Pr(w) \right)$$

The part in brackets is simply our definition of the **Event** "$X = k$". Thus we can rewrite it as

$$= \sum_{\forall\, k} k \cdot \Pr(X = k)$$

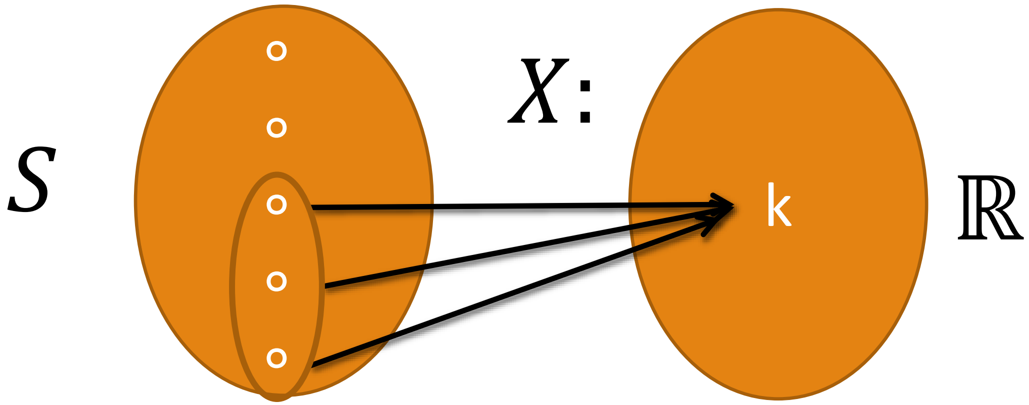
Random Variable: $X: S \to \mathbb{R}$

Event: "$X = k$" $\leftrightarrow \{w \in S: X(w) = k\}$

$k \in$ range of function X.

Instead of looking at every element of S, we look at every **Event** defined by the range of X, that is, all values X can take.

Sum over **Events** instead of **Outcomes**



These **Events** are defined by "$X = k$". So every element in the **Event** is mapped to the same value k.

Likewise, being **Events** we have tools to compute the probability $\Pr(X = k)$.

Given our expected value definition, we gather all values $X(w)$ with the same image to the same **Event**.

$$E(X) = \sum_{w \in S \leftarrow domain\ of\ X} X(w) \cdot \Pr(w)$$

$$E(X) = \sum_{k \in range\ of\ X} k \cdot \Pr(X = k)$$

These expressions are the same (different order).

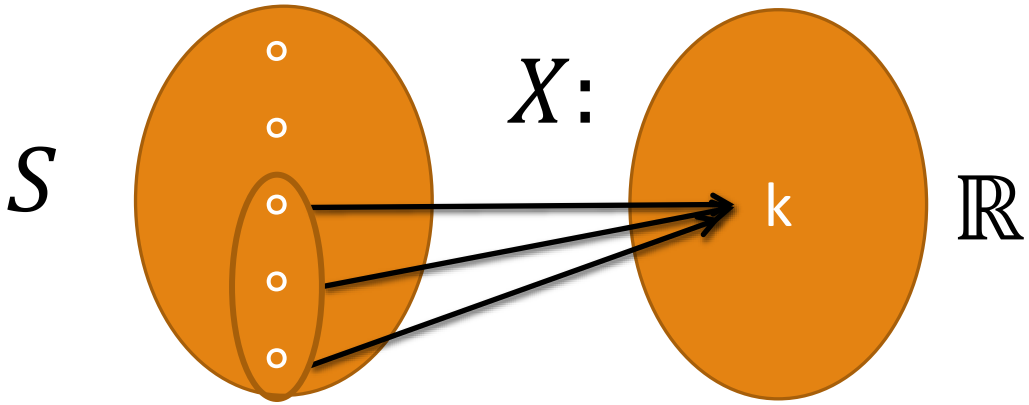
Random Variable: $X: S \to \mathbb{R}$

Event: "$X = k$"$\leftrightarrow \{w \in S: X(w) = k\}$

$k \in$ range of function X.

Instead of looking at every element of S, we look at every **Event** defined by the range of X, that is, all values X can take.

Sum over **Events** instead of **Outcomes**



$$E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)$$

$$E(X) = \sum_{k} k \cdot \Pr(X = k)$$

These expressions are the same.

All we have done to go from one to the other is we changed the order that we summed over the elements $w \in S$.

Next we will look at the 3rd technique, **Linearity of Expectation.**

$$\boxed{E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)}$$

$$E(X) = \sum_{k} k \cdot \Pr(X = k)$$

Linearity of Expectation:

Given two random variables X and Y,

$E(X + Y) = E(X) + E(Y).$

"The expected value of the sum is equal to the sum of the expected values."

We will show this follows directly from the first expression above.

We introduce a third random variable

$$Z = X + Y$$

For any $w \in S$ function $Z(w)$ will give us a value, and that value is exactly the sum of the other two functions:

$$Z(w) = X(w) + Y(w)$$

If we want $E(Z)$ we go with the definition of **Expected Value**:

$$E(Z) = \sum_{w \in S} Z(w) \cdot \Pr(w)$$

$$\boxed{E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)}$$

$$E(X) = \sum_{k} k \cdot \Pr(X = k)$$

Linearity of Expectation:

Given two random variables X and Y,

$E(X + Y) = E(X) + E(Y)$.

"The expected value of the sum is equal to the sum of the expected values."

We will show this follows directly from the first expression above.

$$Z = X + Y$$

$$Z(w) = X(w) + Y(w)$$

$$E(Z) = \sum_{w \in S} Z(w) \cdot \Pr(w)$$

$$= \sum_{w \in S} [X(w) + Y(w)] \cdot \Pr(w)$$

$$= \sum_{w \in S} X(w) \cdot \Pr(w) + \sum_{w \in S} Y(w) \cdot \Pr(w)$$

$$= E(X) + E(Y)$$

$$\boxed{E(X) = \sum_{w \in S} X(w) \cdot \Pr(w)}$$

$$E(X) = \sum_{k} k \cdot \Pr(X = k)$$

Linearity of Expectation:

Given two random variables X and Y,

$E(X + Y) = E(X) + E(Y)$.

"The expected value of the sum is equal to the sum of the expected values."

We will show this follows directly from the first expression above.

This will work for any number of Random Variables:

$$Z(w) = z_1(w) + z_2(w) + \cdots + z_n(w)$$

$$Z(w) = \sum_{i=1}^{n} z_i(w)$$

$$E(Z) = \sum_{i=1}^{n} \sum_{w \in S} z_i(w) \cdot \Pr(w)$$

$$= \sum_{i=1}^{n} E(z_i)$$

Roll fair red die and fair blue die.

$$S = \{\ (i,j) \mid 1 \le i \le 6, 1 \le j \le 6\ \}$$

$X: S \rightarrow\ \mathbb{R}$: $X =$ red + blue: $X(i,j) = i + j$

We know expected value of one die is 3.5, and of 2 dice it is 7, which is twice the value.

This is not a coincidence, and can be verified using linearity of expectation.
Define random variables:

$$red(i,j) = i$$
$$blue(i,j) = j$$

Thus $X = red + blue$

$$E(red) = 3.5$$
$$E(blue) = 3.5$$

Using linearity of expectation:

$$E(X) = E(red + blue) = E(red) + E(blue)$$
$$= \frac{7}{2} + \frac{7}{2} = 7$$

	1	**2**	**3**	**4**	**5**	**6**
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Couple have a child.

Child is a boy, the couple is disappointed since they wanted a girl.

Have another child – another boy.

Keep having children until they have a girl, at which point they stop having children.

[boy probability ½, girl probability ½, independent of the gender of the other children]

If **everyone in the world** does this, will there be more girls than boys in the world, or more boys than girls in the world?

First we will develop a framework to help us solve this problem.

This is an infinite probability space that we will apply Expected Value to.

$0 < p < 1$:
Experiment -> success with prob p
-> failure prob $1 - p$

Instead of children, we will flip a (possibly unfair) coin.

Coin comes up:
H with prob p
T with prob $1 - p$

Flip coins until H, each coin flip is independent.

Define a Random Variable
$X =$ number of flips

What is $E(X)$?

What is the sample space?

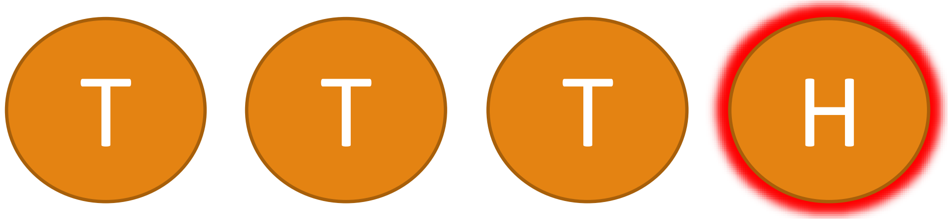
We have seen this before:

$$S = \{H, TH, TTH, TTTH, \dots\}$$

We will define it slightly differently

$$S = \{T^{k-1}H : k \geq 1\}$$

Where k is the number of coin flips. If $k = 1$ then there are 0 tails and 1 heads.



$0 < p < 1$:

Experiment -> success with prob p

-> failure prob $1 - p$

Instead of children, we will flip a (possibly unfair) coin.

Coin comes up:

H with prob p

T with prob $1 - p$

Flip coins until H, each coin flip is independent.

Define a Random Variable

X = number of flips

What is $E(X)$?

$$S = \{T^{k-1}H : k \geq 1\}$$

For any individual outcome of S we can determine the probability. Each coin flip is independent, and thus:

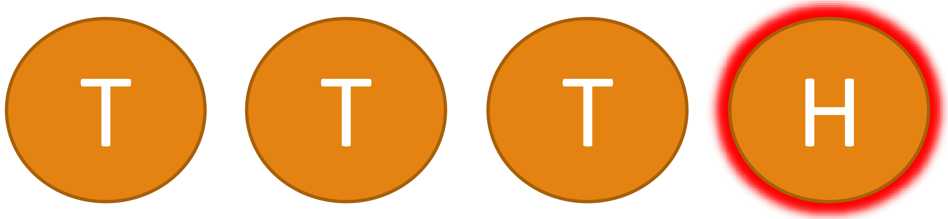
$$\Pr(T^{k-1}H)$$

$$= \Pr(f_1 = T \wedge f_2 = T \wedge \cdots \wedge f_{k-1} = T \wedge f_k = H)$$

$$= \Pr(T) \cdot \Pr(T) \cdot \ldots \cdot \Pr(T) \cdot \Pr(H)$$

$$= \Pr(T)^{k-1} \cdot \Pr(H)$$

$$= (1 - p)^{k-1} \cdot p$$



$$\boxed{\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}}$$

$0 < p < 1$:
Coin comes up:
H with prob p
T with prob $1-p$

Flip coins until H, each coin flip is independent.
$X =$ number of flips

What is $E(X)$?

$$S = \{T^{k-1}H : k \geq 1\}$$

$$\Pr(T^{k-1}H) = (1-p)^{k-1} \cdot p$$

As a sanity check we can verify that the sum of all probabilities of outcomes in S sums to 1.

$$\sum_{k=1}^{\infty} \Pr(T^{k-1}H) = \sum_{k=1}^{\infty} (1-p)^{k-1} \cdot p$$

$$= p \cdot \sum_{k=1}^{\infty} (1-p)^{k-1}$$

$\boxed{\text{Let } i = k-1}$

$$= p \cdot \sum_{i=0}^{\infty} (1-p)^{i}$$

Substitute $(1-p)$ for x:

$$= p \cdot \frac{1}{1-(1-p)} = p \cdot \frac{1}{p} = 1$$

$$\boxed{\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}}$$

$0 < p < 1$:

Coin comes up:

H with prob p

T with prob $1 - p$

Flip coins until H, each coin flip is independent.

X = number of flips

What is $E(X)$?

$$S = \{T^{k-1}H : k \geq 1\}$$

$$\Pr(T^{k-1}H) = (1-p)^{k-1} \cdot p$$

We will use the expression for $E(X)$ where we iterate over the range of X. What is the range of X?

k is the number of flips in the sequence.

We have $k \geq 1$ and $k \to \infty$. Thus:

$$E(X) = \sum_{k=1}^{\infty} k \cdot \Pr(X = k)$$

What is the **Event** $X = k$? It is when there are exactly k coin flips (ending in heads). Thus

$$\text{“}X = k\text{"} = \{T^{k-1}H\}$$

So $\Pr(X = k) = \Pr\left(\{T^{k-1}H\}\right) = \Pr(T^{k-1}H)$

$$\boxed{\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}}$$

$0 < p < 1$:
Coin comes up:
H with prob p
T with prob $1-p$

Flip coins until H, each coin flip is independent.
$X =$ number of flips

What is $E(X)$?

$$S = \{T^{k-1}H : k \geq 1\}$$

$$\Pr(T^{k-1}H) = (1-p)^{k-1} \cdot p$$

$$E(X) = \sum_{k=1}^{\infty} k \cdot \Pr(X = k)$$

Where

$$\Pr(X = k) = \Pr(T^{k-1}H) = (1-p)^{k-1} \cdot p$$

$$E(X) = \sum_{k=1}^{\infty} k \cdot \Pr(X = k)$$

$$= \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1} \cdot p$$

$$= p \cdot \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1}$$

This is an infinite series we have not seen yet.

$$\boxed{\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}}$$

$0 < p < 1$:
Coin comes up:
H with prob p
T with prob $1 - p$

Flip coins until H, each coin flip is independent.
$X =$ number of flips

What is $E(X)$?

$$S = \{T^{k-1}H : k \geq 1\}$$
$$\Pr(T^{k-1}H) = (1-p)^{k-1} \cdot p$$

$$E(X) = \sum_{k=1}^{\infty} k \cdot \Pr(X = k)$$
$$= p \cdot \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1}$$

Without the p we have:

$$1 + 2 \cdot (1-p) + 3 \cdot (1-p)^2 + 4 \cdot (1-p)^3 \ldots$$

If the k were gone we understand how to solve this.

We will look at the general form of this expression and use our favourite infinite sum…

$$\boxed{\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}}$$

The general form of $\sum_{k=1}^{\infty} k \cdot (1-p)^{k-1}$ is:

$$\sum_{k=1}^{\infty} k \cdot x^{k-1}$$

We want to find a closed form. We know this:

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

If we differentiate both sides of this expression with respect to x, they will still be equal. The derivative of the LHS:

$$\frac{d}{dx} \cdot \sum_{k=0}^{\infty} x^k = \sum_{k=0}^{\infty} k \cdot x^{k-1} = \sum_{k=1}^{\infty} k \cdot x^{k-1}$$

And the RHS:

$$\frac{d}{dx} \cdot \frac{1}{1-x} = \frac{0+1}{(1-x)^2} = \frac{1}{(1-x)^2}$$

Thus we have shown that:

$$\sum_{k=1}^{\infty} k \cdot x^{k-1} = \frac{1}{(1-x)^2}$$

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

$$\sum_{k=1}^{\infty} k \cdot x^{k-1} = \frac{1}{(1-x)^2}$$

Coin comes up:
H with prob p
T with prob $1-p$

Flip coins until H, $X =$ number of flips

$$S = \{T^{k-1}H : k \geq 1\}$$

$$\Pr(T^{k-1}H) = (1-p)^{k-1} \cdot p$$

$$E(X) = \sum_{k=1}^{\infty} k \cdot \Pr(X = k)$$

$$= p \cdot \sum_{k=1}^{\infty} k \cdot (1-p)^{k-1}$$

$$= p \cdot \frac{1}{(1-(1-p))^2}$$

$$= p \cdot \frac{1}{p^2}$$

$$= \frac{1}{p}$$

Thus the expected number of trials of an experiment with probability p of success is $\frac{1}{p}$.

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$
$$\sum_{k=1}^{\infty} k \cdot x^{k-1} = \frac{1}{(1-x)^2}$$

Coin comes up:
H with prob p
T with prob $1 - p$

Flip coins until H, $X =$ number of flips

$$S = \{T^{k-1}H : k \geq 1\}$$
$$\Pr(T^{k-1}H) = (1-p)^{k-1} \cdot p$$

$$E(X) = \frac{1}{p}$$

Thus the expected number of trials of an experiment with probability p of success is $\frac{1}{p}$.

So assume that each coin flip lands on heads with probability $\frac{1}{2}$ and tails with probability $\frac{1}{2}$. What is the expected number of times we flip the coin until we see heads?

This is a nice clean result that can be applied in many different places.

A couple keeps having children until they have a girl.

[boy probability ½, girl probability ½, independent of the gender of the other children]

Success: have a girl
Failure: have a boy

How many children do we expect they will have?

Let X be the number of children a couple has.

$$E(X) = \frac{1}{p}$$
$$= \frac{1}{1/2}$$
$$= 2$$

We would expect the couple to have 2 children on average.

A couple keeps having children until they have a girl.

[boy probability ½, girl probability ½, independent of the gender of the other children]

Success: have a girl
Failure: have a boy

Let B be the number of boys born.
Let G be the number of girls born.
Let X be the total number of children a couple has.

$$X = B + G$$

$$E(X) = 2$$
$$E(X) = E(G) + E(B)$$

G is a constant, $G = 1$. Thus:

$$E(G) = 1$$
$$E(B) = 1,$$

If all couples in the world did this, then on average there would be the same number of boys as girls.

What if a couple keeps having children until they have two girls? Does this change the expected number of girls and boys?

As before, with the coin flip game, we can break it down into two rounds.