COMP S264F Unit 8: Conditional Probability, Random Variables

Dr. Keith Lee
School of Science and Technology
The Open University of Hong Kong

Overview

- Conditional probability
- Independent events
  - Independent events E and F: **p(E | F) = p(E)**
  - Two or more independent events
- Not equally likely outcomes
  - Biased coin vs Unbiased coin
- Random variables
- Expected values
  - Sum rule
  - Product rule (*for independent random variables*)

Conditional probability

Suppose we flip a *fair* coin three times.

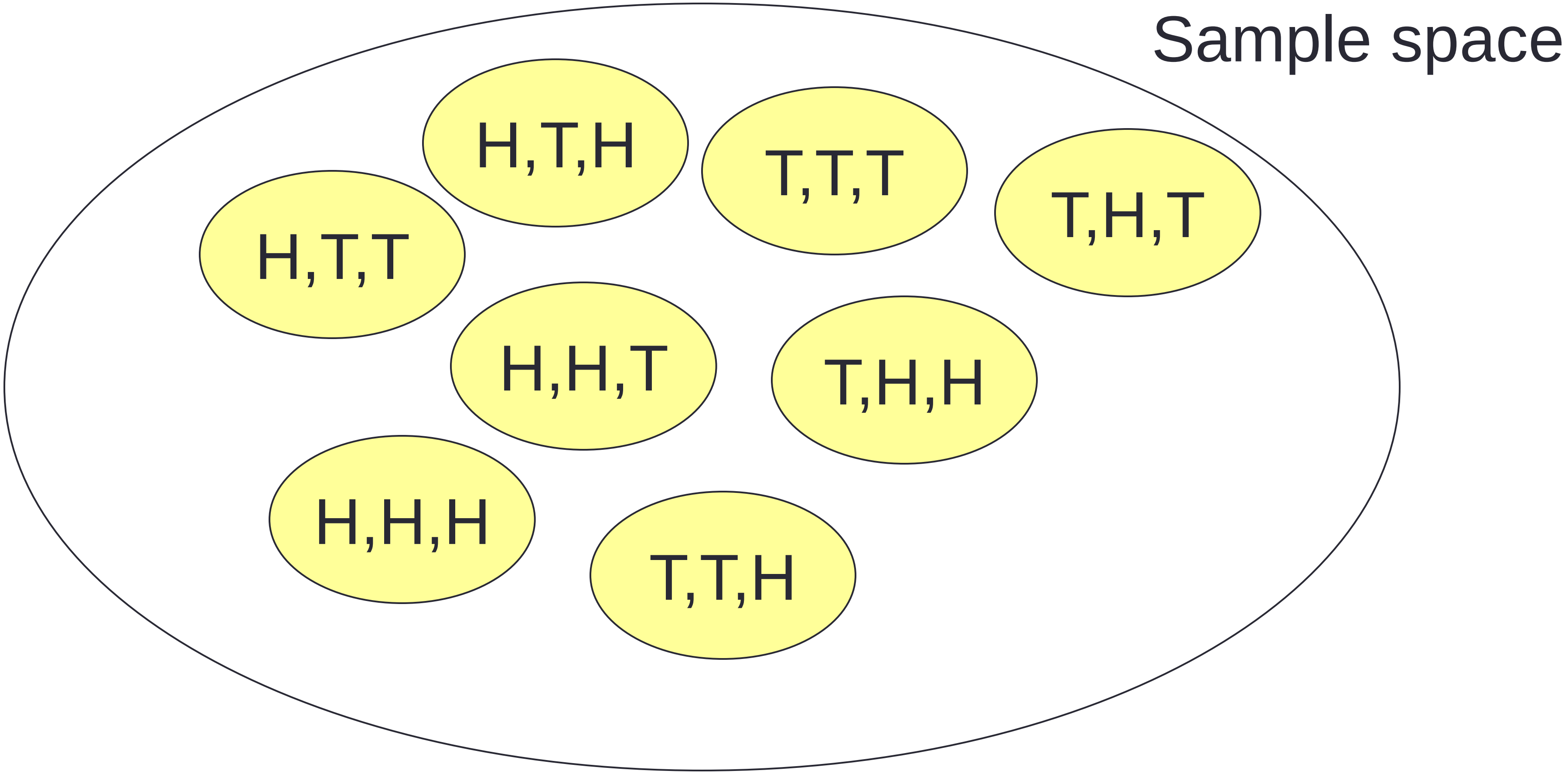
What is the probability that we get 2 or more tails? $\frac{1}{2}$

Again we are interested to determine the probability of this event subject to a certain condition.

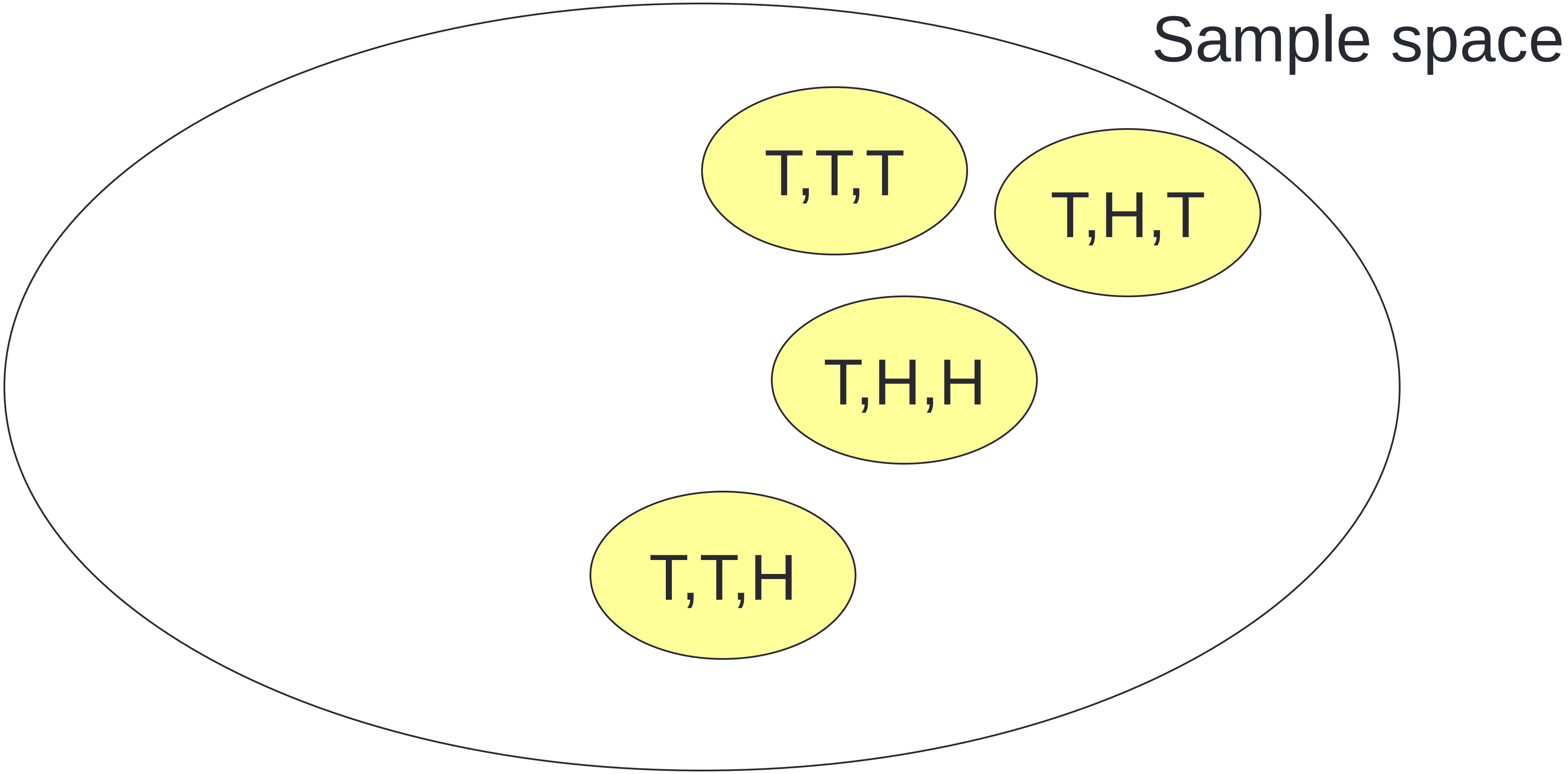
- What is the probability that we get 2 or more tails given that the first flip gives a tail? $\frac{3}{4}$
- What is the probability that we get 2 or more tails given that the first flip gives a head?
- What is the probability that we get 3 tails given that the first two flips give tails?

NB. Knowing that a condition holds may increase or decrease the probability of an event.

Flipping a coin 3 times



Given that the first flip gives a tail



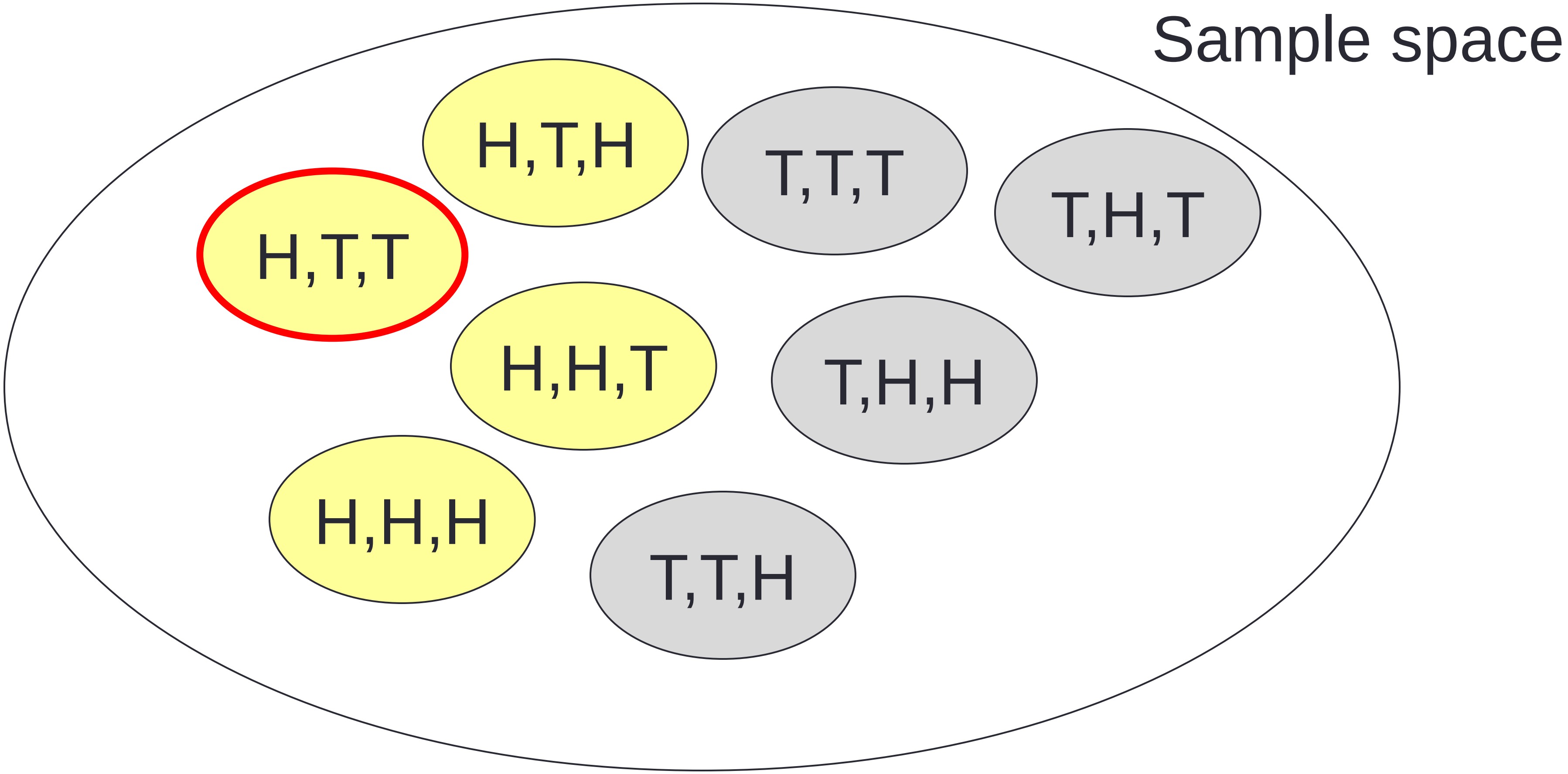
Conditional Probability

- Definition: Let E and F be events w.r.t. (with respect to) a sample space.
- The *conditional probability* of E given F, denoted by $p(E \mid F)$, is $\dfrac{p(E \cap F)}{p(F)}$. (We assume that $p(F) > 0$.)

- E.g., W.r.t. the experiment of flipping a fair coin three times, let **E** be the event that we get 2 or more tails, and let **F** be the event that we get a head in the first flip.

- $p(F) = \frac{1}{2}$; $p(E \cap F) = \frac{1}{8}$; $p(E \mid F) = \dfrac{p(E \cap F)}{p(F)} = \dfrac{1}{8} \cdot \dfrac{2}{1} = \dfrac{1}{4}$

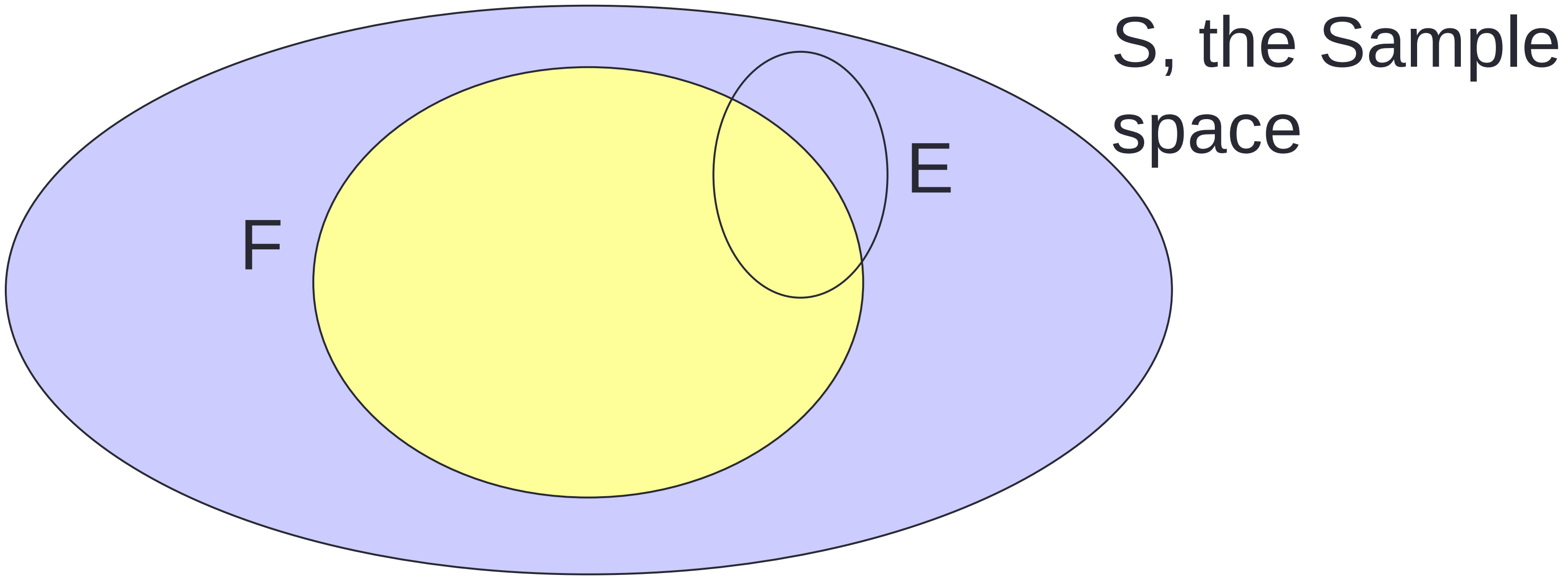
Given that the first flip gives a head



A simple interpretation

(assume every outcome is equally likely)

$$p(E \mid F) = \frac{p(E \cap F)}{p(F)} = \frac{|E \cap F|/|S|}{|F|/|S|} = \frac{|E \cap F|}{|F|}$$



Independence

- Consider an experiment of flipping a fair coin 2 times.
- Let E be the event that the 2nd flip gives a head.
- Let F be the event that the 1st flip gives a head.

- Note that $p(E \mid F) = p(E) = \frac{1}{2}$.

- This is intuitive as the 1st and 2nd flips are conducted "independently".
- Therefore, the fact that F happens does not change the probability that E happens.

Independence: p(E | F) = p(E)

Another interpretation: $p(E \cap F) = p(E) \cdot p(F)$.

$$p(E \mid F) = p(E)$$

$$\Leftrightarrow \quad \frac{p(E \cap F)}{p(F)} = p(E)$$

$$\Leftrightarrow \quad p(E \cap F) = p(E) \cdot p(F)$$

What about "$p(F \mid E) = p(F)$" ?

$$\underbrace{p(F \mid E) = p(F)}_{\text{Knowing E doesn't help}} \;\Leftrightarrow\; p(F \cap E) = p(E) \cdot p(F) \;\Leftrightarrow\; \underbrace{p(E \mid F) = p(E)}_{\text{Knowing F doesn't help}}$$

Independence: Definition

- Two events are said to be **independent** if
  $p(E \mid F) = p(E)$
  (or equivalently, $p(E \cap F) = p(E) \cdot p(F)$).

Example:

- Suppose two fair coins, labeled A and B, are flipped together.
- Let E be the event that coin A comes up a tail.
- Let F be the event that coin B comes up a tail.
- Are E and F independent? Yes.

  (1) $p(E) = p(F) = \frac{1}{2}$; thus, $p(E) \cdot p(F) = \frac{1}{4}$

  (2) $p(E \cap F) = \frac{1}{4}$.

Independence: Example

- Suppose a fair coin is flipped 10 times.
- Let E be the event that there is an *odd* number of tails.
- Let F be the event that the first flip comes up a tail.
- Are E and F independent?
- p(E) = ?

- What is p(E | F) ?

Independence: Example (con't)

- Suppose a fair coin is flipped 10 times.
- Let E be the event that there is an *odd* number of tails.
- Let F be the event that the first flip comes up a tail.
- Are E and F independent?

- Let E_i be the event that there are i tails.
- $|E_1| = C(10, 1) = 10$
- $|E_3| = C(10, 3) = 120$
- $|E_5| = C(10, 5) = 252$
- $|E_7| = C(10, 7) = 120$
- $|E_9| = C(10, 9) = 10$

Total = 512

- $p(E) = \frac{512}{2^{10}} = \frac{512}{1024} = \frac{1}{2}$

Independence: Example (con't)

- Suppose a fair coin is flipped 10 times.
- Let E be the event that there is an *odd* number of tails.
- Let F be the event that the first flip comes up a tail.
- Are E and F independent?

- Let E_i be the event that there are i tails.
- $|E_1 \cap F| = C(9, 0) = 1$
- $|E_3 \cap F| = C(9, 2) = 36$
- $|E_5 \cap F| = C(9, 4) = 126$
- $|E_7 \cap F| = C(9, 6) = 84$
- $|E_9 \cap F| = C(9, 8) = 9$

Total = 256

- $p(E \mid F) = \frac{|E \cap F|}{|F|} = \frac{256}{2^9} = \frac{256}{512} = \frac{1}{2}$.
- As p(E | F) = p(E), E and F are independent.

- Another approach:

Flipping 9 coins:

|even tails| = |even heads|

|even tails| = |odd heads|

(as they are symmetric)

Thus, all of them equal

$\frac{1}{2} \times$ |sample space|

Independence: Examples

- Suppose we flip a fair coin 100 times.

Are the following events independent?

- The number of heads obtained is odd.
- The first 99 flips give all heads.

Are the following events independent?

- The number of heads obtained is odd.
- There are at least 99 heads.

Complement

- Suppose E and F are two independent events.

Complement of F

- Are E and $\overline{F}$ independent? <u>YES</u>. Intuitively, if the fact that F happens does not change the probability of E, then the fact that F doesn't happen does not matter, too.

- Note that $E = (E \cap F) \cup (E \cap \overline{F})$.
- Then,

$$p(E) = p(E \cap F) + p(E \cap \overline{F}) = p(E)\, p(F) + p(E \cap \overline{F}).$$

- Thus,

$$p(E \cap \overline{F}) = p(E)\,(1 - p(F)) = p(E)\, p(\overline{F}).$$

Two or more independent events

- Two or more events E_1, E_2, …, E_n are said to be independent (mutually independent) if

  for all subsets X of {1, 2, …, n},

$$p\Big(\bigcap_{i \in X} E_i\Big) = \prod_{i \in X} p(E_i) \ .$$

- In particular,

$$p(E_1 \cap E_2 \cap \cdots \cap E_n) = p(E_1)p(E_2)\dots p(E_n).$$

Not equally likely Outcomes

- Consider a sample space $S = \{x_1, x_2, \ldots, x_n\}$.
- Suppose that the outcomes may not be equally likely; let $p(x_i)$ denote the probability that x_i occurs.
- Of course, we require that $p(x_1) + p(x_2) + \ldots + p(x_n) = 1$.

- Consider an event $E = \{x_{i1}, x_{i2}, \ldots, x_{im}\} \subseteq S$.
- The probability of E, denoted by p(E), is

  $p(x_{i1}) + p(x_{i2}) + \ldots + p(x_{im})$.

- NB. *If the outcomes are equally likely,* $p(E) = \dfrac{m}{n}$.

Not equally likely Outcomes: Example

- Consider a biased dice.
- $p(1) = p(2) = 0.2$;
- $p(3) = p(4) = 0.25$;
- $p(5) = p(6) = 0.05$.

Not equally likely Outcomes

- $p(small) = p(1) + p(2) + p(3) = 0.2 + 0.2 + 0.25 = 0.65$
- $p(even) = p(2) + p(4) + p(6) = 0.2 + 0.25 + 0.05 = 0.5$

Question: Biased versus unbiased coins

- Suppose you are given a coin for which the probability of HEAD is $p \leq \frac{1}{4}$. But the exact probability is not known.
- How can you use this coin to generate unbiased coin-flips?

Solution:

- We need to get 2 cases of equal probability.
- Flip the coin twice:
  - $p(HT) = p\,(1-p)$
  - $p(TH) = (1-p)\,p = p\,(1-p)$
- Thus, we can generate unbiased coin-flips by two coin flips:
  - If the result is HT, then treat it as a HEAD.
  - If the result is TH, then treat it as a TAIL.
  - Otherwise, if the result is HH or TT, repeat again.

Random Variables

- In many cases, we associate a *numeric value* with each outcome of an experiment.

For instance,
consider the experiment of flipping a coin 10 times,
each outcome can be associated with

- the number of heads,
- the number of tails,
- the difference between heads & tails, etc.

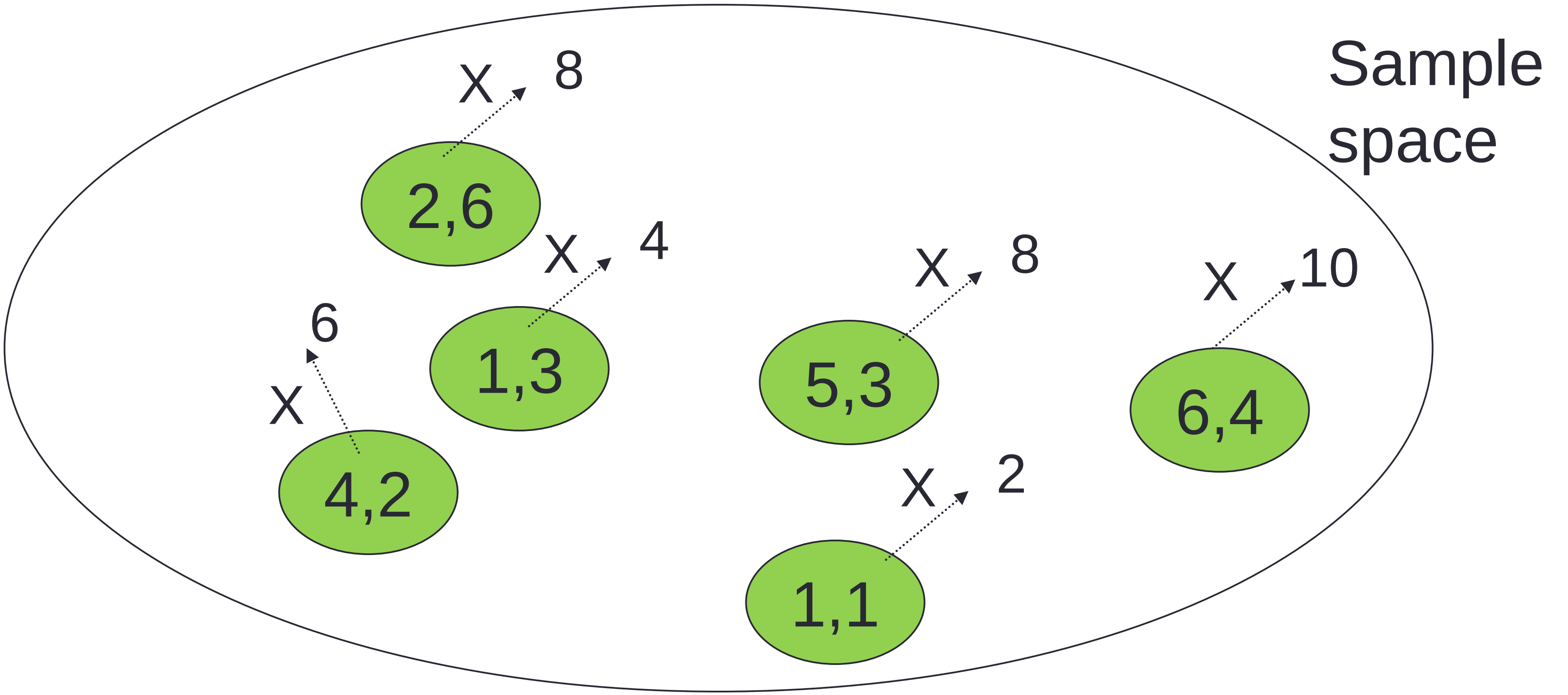
- All the quantities above are called ***random variables***.

Random Variables: Definition

- With respect to an experiment,
a random variable is a ***<u>function</u>***
  - from the set of possible outcomes
  - to the set of real numbers

- NB. A random variable is characterized by the sample space of an experiment.

- **Example:** Let X be the sum of the numbers that appear when a pair of dice is rolled.
- There are 36 possible outcomes, each defines a value of X (within the range from 2 to 12).



Notation: X((5,4)) = 9, or for the outcome (5,4), X = 9.

Random variables and events

A more intuitive way to look at random variable X is to examine the probability of each possible value of X.

E.g., consider the previous example:

- Let p(*X=3*) be the *probability of the event* that the sum of the two dice is 3.
  NB. This event comprises two outcomes, (1,2) and (2,1).
- Note that $p(X=3) = \frac{2}{36}$.

- In general, for any random variable ***v***, "***v*** = *i* " defines an event, and p(***v*** = *i*) = the sum of the probability of all the outcomes *y* such that ***v****(y)* = i.

- $\sum_{i \in \text{the range of } v} p(v = i) = 1$

Expected value

In the previous example, what is the expected value (average value) of X?

Expected value of X

$$= (2 + 3\times2 + 4\times3 + 5\times4 + 6\times5 + 7\times6 + 8\times5 + 9\times4 + 10\times3 + 11\times2 + 12) / 36$$

$$= \frac{252}{36} = 7$$

Out of the 36 outcomes,

(1,1): X=2

(1,2), (2,1): X=3

(1,3), (2,2), (3,1): X=4

(1,4), (2,3), (3,2), (4,1): X=5

(1,5), (2,4), (3,3), (4,2), (5,1): X=6

(1,6), (2,5), (3,4), (4,3), (5,2), (6,1): X=7

(2,6), (3,5), (4,4), (5,3), (6,2): X=8

(3,6), (4,5), (5,4), (6,3): X=9

(4,6), (5,5), (6,4): X=10

(5,6), (6,5): X=11

(6,6): X=12

Expected value: Definition

- Consider an experiment with sample space **S**. For any outcome *y* in **S**, let p(*y*) be the probability y occurs.

X: **S** $\rightarrow \mathbb{R}$ (i.e., Real)

- Let X be a random variable. That is, every outcome *y* in **S** defines a value of X, denoted by X(*y*).
- We define the expected value of X, denoted by E(X), to be

$$\sum_{y \in \mathbf{S}} p(y) \cdot X(y)$$

or equivalently,

$$\sum_{i \,\in\, the\ range\ of\ X} p(X = i) \cdot i$$

Expected value: Example

What is the *expected number of heads* in flipping a fair coin four times?

Prob[# of heads = 1]

$$
\begin{aligned}
&1 \times \boxed{C(4,1)\ 1/2\ (1/2)^3} \\
&+ 2 \times \underline{C(4,2)\ (1/2)^2\ (1/2)^2} \\
&+ 3 \times \underline{C(4,3)\ (1/2)^3\ 1/2} \\
&+ 4 \times \underline{(1/2)^4}
\end{aligned}
$$

$$= \frac{4}{16} + 2 \times \frac{6}{16} + 3 \times \frac{4}{16} + 4 \times \frac{1}{16}$$

$$= 2$$

Example: Network Protocol

```
Repeat
   flip a fair coin twice;
   if "head + head", then send a packet to the network;
until "sent"
```

- What is the expected number of iterations used by the protocol?

Example: Network Protocol (cont')

- Let p be the probability of success within each trial.
- Let q be the probability of failure within each trial.
- Expected value

$$= \sum_{i=1 \text{ to } \infty} (\text{prob. of sending the packet in the } i\text{-th trial}) \cdot i$$

$$= \sum_{i=1 \text{ to } \infty} (pq^{i-1}) \cdot i$$

$$= p \cdot \sum_{i=1 \text{ to } \infty} q^{i-1} \cdot i$$

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

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

Example: String comparisons

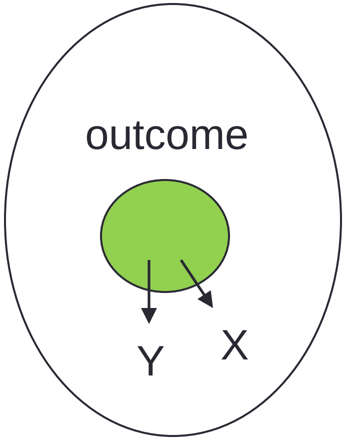
- Let L be a linked list with each node containing a name.
- Suppose we want to search L for a given name X.
- The number of string comparisons involved ranges from 1 to n, where n is the number of nodes in L.

- What is the expected number of string comparisons?
- This question is not well-defined! # of string comparisons is a random variable but w.r.t. what sample space?
- Let us assume that X has probability 0.5 to appear in L, and matches each node with equal probability.
- Then, the expected number of comparisons

$$
\begin{aligned}
&= \frac{1}{2} \cdot n + \frac{1}{2n} \cdot (1 + 2 + \cdots + n) \\
&= \frac{n}{2} + \frac{1}{2n} \cdot \frac{n(n+1)}{2} = \frac{2n + (n+1)}{4} = \frac{3n+1}{4}
\end{aligned}
$$

Useful rules for deriving expected values
Sum rule

- Let X and Y be random variables on a space S.
- Then, X + Y is also a random variable on S, and
- E(X + Y) = E(X) + E(Y).



Proof:

$$
\begin{aligned}
{\color{red}E(X+Y)} &= \sum_{t \in S} p(t) \cdot (X(t) + Y(t)) \\
&= \sum_{t \in S} p(t) \cdot X(t) + \sum_{t \in S} p(t) \cdot Y(t) \\
&= E(X) + E(Y)
\end{aligned}
$$

Sum rule: Example

- Use the "sum rule" to derive the expected value of the sum of the numbers when we roll a pair of fair dice (denoted by X).
- Suppose the dice are colored red & blue.
- Let X_1 be the number on the red dice when we roll a pair of red & blue dice, and similarly X_2 for the blue dice.

  $E(X_1) = E(X_2) = ?$

- Obviously, $X = X_1 + X_2$.
- Thus, $E(X) = E(X_1 + X_2) = ?$

What about product?

- Let X and Y be two random variables of a space S.
- Is E(XY) = E(X) E(Y) ?

- **Example 1:** Consider tossing a coin twice. Associate “head” with 2 and “tail” with 1.
- What is the expected value of the product of the numbers obtained in tossing a coin twice.
  - $(1,1) \rightarrow 1$; $(1,2) \rightarrow 2$; $(2,1) \rightarrow 2$; $(2,2) \rightarrow 4$
  - Expected product = $\frac{(1+2+2+4)}{4} = 2.25$
  - Expected value of 1st flip = $\frac{(1+2)}{2} = 1.5$
  - Expected value of 2nd flip = $\frac{(1+2)}{2} = 1.5$
- Note that $2.25 = 1.5 \times 1.5$!

Counterexample

- Consider the previous experiment again.
- Define a random variable X as follows:
  X = (the first number) $\times$ (the sum of the two numbers)
- $(1,1) \rightarrow 1 \times 2 = 2$
- $(1,2) \rightarrow 1 \times 3 = 3$
- $(2,1) \rightarrow 2 \times 3 = 6$
- $(2,2) \rightarrow 2 \times 4 = 8$
- Expected value of 1st number = 1.5
- Expected sum = $\dfrac{(2+3+3+4)}{4} = 3$
- NB. $4.75 \neq 1.5 \times 3$

Expected value = 4.75

Why? Because the first # & the sum are not independent.

Independent random variables

- Two random variables X and Y on a sample space S are independent if, for all real numbers r_1 and r_2,

$$p(X = r_1 \text{ and } Y = r_2) = p(X = r_1) \times p(Y = r_2) .$$

- Alternative definition:
  X and Y are independent if for all real numbers r_1 and r_2, the events "$X = r_1$" and "$Y = r_2$" are independent.

Product rule: If X and Y are independent random variables on a space S, then E(XY) = E(X) E(Y).

Proof:

$$\begin{aligned}
E(XY) &= \sum_{W \in \mathbb{R}} W \cdot p(XY = W) \\
&= \sum_{W \in \mathbb{R}} \sum_{a,b \in \mathbb{R} \text{ such that } ab=W} W \cdot p(X = a \text{ and } Y = b) \\
&= \sum_{a,b \in \mathbb{R}} (ab) \cdot p(X = a \text{ and } Y = b) \\
&= \sum_{a,b \in \mathbb{R}} (ab) \cdot p(X = a) \cdot p(Y = b) \quad \text{(as } X \text{ \& } Y \text{ are independent)} \\
&= \sum_{a,b \in \mathbb{R}} (a \cdot p(X = a)) \cdot (b \cdot p(Y = b)) \\
&= \sum_{a \in \mathbb{R}} \left[a \cdot p(X = a) \cdot \sum_{b \in \mathbb{R}} b \cdot p(Y = b) \right] \\
&= \left[\sum_{a \in \mathbb{R}} a \cdot p(X = a) \right] \cdot \left[\sum_{b \in \mathbb{R}} b \cdot p(Y = b) \right] = E(X)E(Y)
\end{aligned}$$