Probability Review

Random Variable
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 difference between heads & tails, etc.

Each quantity above is called a *random variable*; note that its value depends on the outcome of the experiment and is not fixed in advance.
Formally speaking

With respect to an experiment, a random variable is a \textit{function}:

\begin{itemize}
  \item from the set of possible outcomes $\Omega$
  \item to the set of real numbers.
\end{itemize}

NB. A random variable is characterised by the sample space of an experiment.
Example: Let $X$ be the sum of the numbers obtained by rolling a pair of fair dice. There are 36 possible outcomes, each defines a value of $X$ (in the range from 2 to 12).
Random variables and events

A more intuitive way to look at the 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. This event comprises two outcomes, (1,2) and (2,1).
- $p(X=3) = 2/36$. 

Random variables and events

In general, for any random variable $X$, 
$p(X = i) = \text{the sum of the probability of all the outcomes } y \text{ such that } X(y) = i.
\sum_{i \in \text{the range of } X} p(X = i) = 1$
Expected value

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

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 of $X = \frac{2 + 3 \times 2 + 4 \times 3 + 5 \times 4 + 6 \times 5 + 7 \times 6 + 8 \times 5 + 9 \times 4 + 10 \times 3 + 11 \times 2 + 12}{36}$

$= 7$
Consider an experiment with a sample space $S$. For any outcome $y$ in $S$, let $p(y)$ be the probability $y$ occurs.

Let $X$ be an integer random variable over $S$. That is, every outcome $y$ in $S$ defines a value of $X$, denoted by $X(y)$.

We define the expected value of $X$ to be

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

or equivalently,

$$\sum_{i \in Z} p(X = i) i$$
Example

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

\[
\begin{align*}
1 \times \binom{4}{1} \frac{1}{2} \left(\frac{1}{2}\right)^3 \\
+ 2 \times \binom{4}{2} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 \\
+ 3 \times \binom{4}{3} \left(\frac{1}{2}\right)^3 \frac{1}{2} \\
+ 4 \times \left(\frac{1}{2}\right)^4
\end{align*}
\]

= \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

Let $p$ be the probability of success within each trial. Let $q$ be the probability of failure within each trial.

$$\sum_{i=1}^{\infty} \left( \text{prob of sending the packet in the i-th trial} \right) i$$

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

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

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

$$= \frac{1}{p}$$
Useful rules for deriving expected values

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.**

\[
E(X + Y) = \sum_{t \in S} p(t) (X(t) + Y(t))
\]

\[
= \sum_{t \in S} p(t) X(t) + \sum_{t \in S} p(t) Y(t)
\]

\[
= E(X) + E(Y)
\]
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, $E(X) = E(X_1) + E(X_2) = \ ?$
What about product?

Let $X$ and $Y$ be two random variables of a space $S$.

Is $E(XY) = E(X) E(Y)$?
What about product?

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

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 $= (1+2+2+4) / 4 = 2.25$
- Expected value of 1st flip: $(1+2) / 2 = 1.5$
- Expected value of 2nd flip: $(1+2) / 2 = 1.5$

Note that $2.25 = 1.5 \times 1.5$!
Consider the previous experiment again. Define a random variable $X$ as follows:

$X = (\text{the first number}) \times (\text{the sum of the two numbers})$

- $(1,1) \rightarrow 1 \times 2 = 2$
- $(1,2) \rightarrow 1 \times 3 = 3$
- $(2,1) \rightarrow 2 \times 3 = 6$
- $(2,2) \rightarrow 2 \times 4 = 8$

- Expected value of 1st number = 1.5
- Expected sum = $(2 + 3 + 3 + 4)/4 = 3$

Expected value = 4.75

$4.75 \neq 1.5 \times 3$

Why! Because the first # & the sum are not independent.
Counter Example

Consider the experiment of flipping a fair coin twice.

Expected (number of) heads = 1
Expected tails = 1
Expected heads \times expected tails = 1

Expected (heads \times tails) = 2 / 4 = 0.5
Independent random variables

Two random variables $X$ and $Y$ over a sample space $S$ are independent if

for all real numbers $r_1$ and $r_2$,
the events “$X = r_1$” and “$Y = r_2$” are independent,

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


Product rule.

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