Random Variables

Questions about outcomes ...

Experiment: roll two dice.
Sample Space: \{(1,1),(1,2), \ldots ,(6,6)\} = \{1, \ldots ,6\}^2
How many pips?

Experiment: flip 100 coins.
Sample Space: \{\text{HHH} \ldots \text{H}, \text{THH} \ldots \text{H}, \ldots , \text{TTT} \ldots \text{T}\}
How many heads in 100 coin tosses?

Experiment: choose a random student in cs70.
Sample Space: \{Adam, Jin, Bing ..., Angeline\}
What midterm score?

Experiment: hand back assignments to 3 students at random.
Sample Space: \{123, 132, 213, 231, 312, 321\}
How many students get back their own assignment?

In each scenario, each outcome gives a number.
The number is a (known) function of the outcome.


- **Sample Space**: Set of outcomes, \(\Omega\).
- **Probability**: \(Pr[\omega]\) for all \(\omega \in \Omega\).
  - \(0 \leq Pr[\omega] \leq 1\).
  - \(\sum_{\omega \in \Omega} Pr[\omega] = 1\).
- **Event**: \(A \subseteq \Omega\). \(Pr[A] = \sum_{\omega \in A} Pr[\omega]\).
  - Inclusion/Exclusion: \(Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]\).
  - Simple Total Probability: \(Pr[B] = Pr[A \cap B] + Pr[\overline{A} \cap B]\).
  - Complement: \(Pr[\overline{A}] = 1 - Pr[A]\).
  - Union Bound. Total Probability.
  - **Conditional Probability**: \(Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}\).
  - **Bayes’ Rule**: \(Pr[A_i|B] = \frac{Pr[A_i]Pr[B|A_i]}{\sum_{i} Pr[A_i]Pr[B|A_i]}\).
  - **Product Rule**: \(Pr[A_1 \cap \cdots \cap A_n] = Pr[A_1]Pr[A_2|A_1] \cdots Pr[A_n|A_1 \cap \cdots \cap A_{n-1}]\).
  - **Total Probability/Product**: \(Pr[B] = Pr[B|A]Pr[A] + Pr[B|\overline{A}]Pr[\overline{A}]\).

Random Variables

A random variable, \(X\), for an experiment with sample space \(\Omega\) is a function \(X : \Omega \rightarrow \mathbb{R}\).
Thus, \(X(\cdot)\) assigns a real number \(X(\omega)\) to each \(\omega \in \Omega\).

Function \(X(\cdot)\) defined on outcomes \(\Omega\).
Function \(X(\cdot)\) is not random, not a variable!
What varies at random (among experiments)? The outcome!
Note: Random variable induces partition:
\(A_y = \{\omega \in \Omega : X(\omega) = y\} = X^{-1}(y)\)

Example 1 of Random Variable

Experiment: roll two dice.
Sample Space: \{(1,1),(1,2), \ldots ,(6,6)\} = \{1, \ldots ,6\}^2
Random Variable \(X\): number of pips.
\(X(1,1) = 2\)
\(X(1,2) = 3\),
\(\vdots\)
\(X(6,6) = 12\)
\(X(a,b) = a+b, (a,b) \in \Omega\).
Example 2 of Random Variable

Experiment: flip three coins
Sample Space: \{HHH, THH, HTH, TTH, HHT, THT, HTT, TTT\}
Winnings: if win 1 on heads, lose 1 on tails. \(X\)
\[
X(\text{HHH}) = 3 \quad X(\text{THH}) = 1 \quad X(\text{HTH}) = 1 \quad X(\text{TTH}) = -1
\]
\[
X(\text{HHT}) = 1 \quad X(\text{THT}) = -1 \quad X(\text{HTT}) = -1 \quad X(\text{TTT}) = -3
\]

Handing back assignments

Experiment: hand back assignments to 3 students at random.
Sample Space: \(\Omega = \{123, 132, 213, 231, 312, 321\}\)
How many students get back their own assignment?
Random Variable: values of \(X(\omega)\): \{3, 1, 1, 0, 0, 1\}
Distribution:
\[
X = \begin{cases} 0, \text{ w.p. } 1/3 \\ 1, \text{ w.p. } 1/2 \\ 3, \text{ w.p. } 1/6 \end{cases}
\]

Number of pips in two dice.

“What is the likelihood of getting \(n\) pips?”
\[
\Pr[X = 10] = \frac{3}{36} = \Pr[X^{-1}(10)]; \Pr[X = 8] = \frac{5}{36} = \Pr[X^{-1}(8)].
\]

Distribution

The probability of \(X\) taking on a value \(a\).
Definition: The distribution of a random variable \(X\), is \(\{(a, \Pr[X = a]) : a \in A\}\), where \(A\) is the range of \(X\).
\[
\Pr[X = a] := \Pr[X^{-1}(a)] \text{ where } X^{-1}(a) := \{\omega | X(\omega) = a\}.
\]

Flip three coins

Experiment: flip three coins
Sample Space: \{HHH, THH, HTH, TTH, HHT, THT, HTT, TTT\}
Winnings: if win 1 on heads, lose 1 on tails. \(X\)
\[
X(\text{HHH}) = 3 \quad X(\text{THH}) = 1 \quad X(\text{HTH}) = 1 \quad X(\text{TTH}) = -1
\]
\[
X(\text{HHT}) = 1 \quad X(\text{THT}) = -1 \quad X(\text{HTT}) = -1 \quad X(\text{TTT}) = -3
\]

Number of pips.

Experiment: roll two dice.
Sample Space: \{1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}
Winnings: if win 1 on heads, lose 1 on tails. \(X\)
\[
X = \begin{cases} -3, \text{ w. p. } 1/8 \\ -1, \text{ w. p. } 3/8 \\ 1, \text{ w. p. } 3/8 \\ 3, \text{ w. p. } 1/8 \end{cases}
\]
Expectation - Definition

Definition: The expected value of a random variable $X$ is

$$E[X] = \sum_a a \times \Pr[X = a].$$

The expected value is also called the mean.

According to our intuition, we expect that if we repeat an experiment a large number $N$ times and if $X_1, \ldots, X_N$ are the successive values of the random variable, then

$$\frac{X_1 + \ldots + X_N}{N} \approx E[X].$$

That is indeed the case, in the same way that the fraction of times $X = x$ approaches $\Pr[X = x]$. This (nontrivial) result is called the Law of Large Numbers.

The subjectivist (bayesian) interpretation of $E[X]$ is less obvious.

Expectation and Average

There are $n$ students in the class;

$$X(m) = \text{score of student } m, \text{ for } m = 1, 2, \ldots, n.$$ 

“Average score” of the $n$ students: add scores and divide by $n$:

$$\text{Average} = \frac{X(1) + X(1) + \ldots + X(n)}{n}.$$

Expectation: a useful fact.

Experiment: choose a student uniformly at random. Uniform sample space: $\Omega = \{1, 2, \ldots, n\}$, $\Pr[\omega] = 1/n$ for all $\omega$.

Random Variable: midterm score: $X(\omega)$.

Expectation:

$$E(X) = \sum_{\omega} X(\omega) \Pr[\omega] = \sum_{\omega} X(\omega) \frac{1}{n}.$$

Hence,

$$\text{Average} = \frac{1}{n} \sum_{\omega} X(\omega).$$

This holds for a uniform probability space.

Named Distributions

Some distributions come up over and over again.

...like “choose” or “stars and bars”....

Let’s cover some.
The binomial distribution.

Flip \( n \) coins with heads probability \( p \).
Random variable: number of heads.

Binomial Distribution: \( \Pr[X = i] \), for each \( i \).

How many sample points in event \( *X = i* \)? \( i \) heads out of \( n \) coin flips \( \implies \binom{n}{i} \).

What is the probability of \( \omega \) if \( \omega \) has \( i \) heads? Probability of heads in any position is \( p \).
Probability of tails in any position is \( 1 - p \).
So, we get \( \Pr[\omega] = p^i(1 - p)^{n-i} \).

Probability of \( *X = i* \) is sum of \( \Pr[\omega], \omega \in *X = i* \).

\[
\Pr[X = i] = \binom{n}{i} p^i (1-p)^{n-i}, i = 0,1,\ldots,n : B(n,p) \text{ distribution}
\]

Expectation of Binomial Distribution

Parameter \( p \) and \( n \). What is expectation?

\[
\Pr[X = i] = \binom{n}{i} p^i (1-p)^{n-i}, i = 0,1,\ldots,n : B(n,p) \text{ distribution}
\]

\[
E[X] = \sum_i i \times \Pr[X = i].
\]

Uh oh? Well... It is \( np \).

Proof? After linearity of expectation this is easy.

Waiting is good.

The binomial distribution.

1. \( T \ T \ T \ T \ T \ T \ T \ T \ T \)
2. \( 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \ 1 \)
3. \( m \ 1 \ m \ 1 \ m \ 1 \ m \ 1 \ m \)

(\( m \) times \( (1-p) \) \( p^m \) \((1-p)^{n-m} \))

(\( \binom{n}{m} \) outcomes with \( m \) Hs and \( n-m \) Ts

\[
\Pr[X = m] = \binom{n}{m} p^m (1-p)^{n-m}
\]

Uniform Distribution

Roll a six-sided balanced die. Let \( X \) be the number of pips (dots).
Then \( X \) is equally likely to take any of the values \( \{1,2,\ldots,6\} \). We say that \( X \) is uniformly distributed in \( \{1,2,\ldots,6\} \).

More generally, we say that \( X \) is uniformly distributed in \( \{1,2,\ldots,n\} \) if
\[
\Pr[X = m] = \frac{1}{n} \text{ for } m = 1,2,\ldots,n.
\]

In that case,
\[
E[X] = \sum_{m=1}^n m \Pr[X = m] = \sum_{m=1}^n m \times \frac{1}{n} = \frac{n(n+1)}{2} = \frac{n+1}{2}.
\]

Error channel and...

A packet is corrupted with probability \( p \).
Send \( n + 2k \) packets.
Probability of at most \( k \) corruptions.
\[
\sum_{i=0}^k \binom{n+2k}{i} p^i (1-p)^{n+2k-i}.
\]

Also distribution in polling, experiments, etc.

Geometric Distribution

Let's flip a coin with \( \Pr[H] = p \) until we get \( H \).

For instance:
- \( \omega_1 = H \), or
- \( \omega_2 = T \ H \), or
- \( \omega_3 = T \ T \ T \), or
- \( \omega_{6n} = T \ T \ T \ T \ T \ T \ H \).

Note that \( \Omega = \{ \omega_n, n = 1,2,\ldots \} \).
Let \( X \) be the number of flips until the first \( H \). Then, \( X(\omega_n) = n \).
Also,
\[
\Pr[X = n] = (1-p)^{n-1} p, n \geq 1.
\]
Geometric Distribution

\[ \Pr(X = n) = (1 - p)^{n-1}p, \quad n \geq 1. \]

Hence,\[ \sum_{n=1}^{\infty} \Pr(X = n) = \sum_{n=1}^{\infty} (1 - p)^{n-1}p = p \sum_{n=0}^{\infty} (1 - p)^n. \]

Now, if \( |a| < 1 \), then \( S := \sum_{n=0}^{\infty} a^n = \frac{1}{1-a} \). Indeed, \[ aS = 1 + a + a^2 + a^3 + \cdots \]
\[ (1-a)S = 1 + a + a^2 + a^3 + \cdots - 1 = a + a^2 + a^3 + \cdots = S. \]

Hence, \[ \sum_{n=1}^{\infty} \Pr(X = n) = p \frac{1}{1-(1-p)} = 1. \]

Geometric Distribution

\[ \Pr(X = n) = (1 - p)^{n-1}p, \quad n \geq 1. \]

Note that \[ \sum_{n=1}^{\infty} \Pr(X = n) = \sum_{n=1}^{\infty} (1 - p)^{n-1}p = p \sum_{n=0}^{\infty} (1 - p)^n. \]

One has \[ E[X] = \sum_{n=1}^{\infty} n \Pr(X = n) = \sum_{n=1}^{\infty} n(1 - p)^{n-1}p. \]

Thus, \[ E[X] = p + 2(1-p)p + 3(1-p)^2p + 4(1-p)^3p + \cdots \]
\[ (1-p)E[X] = (1-p)p + 2(1-p)^2p + 3(1-p)^3p + \cdots \]
\[ pE[X] = p + (1-p)p + (1-p)^2p + (1-p)^3p + \cdots \]
by subtracting the previous two identities
\[ = \sum_{n=1}^{\infty} \Pr(X = n) = 1. \]

Hence, \[ E[X] = \frac{1}{p}. \]

Summary

- A random variable \( X \) is a function \( X : \Omega \to \mathbb{R} \).
- \( \Pr(X = a) := \Pr[X^{-1}(a)] = \Pr\{\omega | X(\omega) = a\} \).
- \( \Pr(X \in A) := \Pr[X^{-1}(A)] \).
- The distribution of \( X \) is the list of possible values and their probability: \( \{(a, \Pr[X = a]), a \in \mathbb{R}\} \).
- \( E[X] := \sum_a a \Pr[X = a] \).
- Expectation is Linear.
- \( B(n, p), U[1 : n], G(p), P(\lambda) \).