

Random Variables

Definition. A **random variable** is a quantity whose value depends on chance

Definition. A **random variable** is a function from sample space to \mathbb{R}

Example

Suppose we have sample space $\Omega = \{a, b, c, d, e\}$ and $\mathbb{P}(a) = \mathbb{P}(b) = \mathbb{P}(c) = \mathbb{P}(d) = \mathbb{P}(e) = \frac{1}{5}$. Suppose we also have $X : \Omega \rightarrow \mathbb{R}$ such that $X(a) = 1, X(b) = 2, X(c) = 3, X(d) = 4, X(e) = 5$. X is a random variable.

Example

We toss a coin and roll a die and use their results as our sample space. What random variables could we have?

	H	T
1	(H,1)	(T,1)
2	(H,2)	(T,2)
3	(H,3)	(T,3)
4	(H,4)	(T,4)
5	(H,5)	(T,5)
6	(H,6)	(T,6)

Definition.

$$\mathbb{P}(X = x) = \mathbb{P}(\{\omega \in \Omega | X(\omega) = x\})$$

We say ‘the probability X takes the value x ’. It is standard to use capital letters to denote the random variable, and lower-case letters for the value it takes.

Example

$$\begin{aligned} \mathbb{P}(X = 3) &= \mathbb{P}(\{\omega \in \Omega | X(\omega) = 3\}) \\ &= \mathbb{P}(\{c\}) \\ &= \frac{1}{5} \end{aligned}$$

General Discrete Distribution

Definition. The **probability distribution** of a random variable is a mapping of possible values of the variable and the corresponding probabilities.

Since a random variable must take a value,

$$\sum_x \mathbb{P}(X = x) = 1 \text{ where the sum is over all possible values } X \text{ can take}$$

Example (A (not very) random variable)

$$\mathbb{P}(X = 1) = 1$$

Example (Discrete Uniform)

$$\mathbb{P}(X = x) = \frac{1}{6} \text{ for } 1 \leq x \leq 6$$

Example

$$\mathbb{P}(X = x) = cx(x + 1) \text{ for } 1 \leq x \leq 10. \text{ What is } c?$$

Example (Infinite discrete distribution)

$$\mathbb{P}(X = x) = \frac{1}{2^x}, x = 1, 2, 3, \dots$$

Independence

Definition. We say random variables X and Y are **independent** iff

$$\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$$

for all $x, y \in \mathbb{R}$

Tip

In other words, the **events** $\{\omega \in \Omega | X(\omega) = x\}$ and $\{\omega \in \Omega | Y(\omega) = y\}$ are independent

Example

X and Y are discrete random variables, with distributions

x	1	2	3
$P(X = x)$	0.2	0.5	0.3

y	0	1
$P(Y = y)$	0.4	0.6

- (a) Find $P(X = 2 \text{ and } Y = 0)$.
- (b) Find $P(X + Y = 3)$.

Example

The joint probability table for two discrete random variables S and T is shown below:

	$T = 1$	$T = 2$
$S = 1$	0.12	0.18
$S = 2$	0.28	0.42

- Find the marginal distributions of S and T (i.e., find $P(S = s)$ and $P(T = t)$ for all possible values).
- Determine whether S and T are independent. Justify your answer.

Expectation and Variance

Definition. The **expectation** (mean) of a random variable is

$$\mu = \mathbb{E}[X] = \sum_x x\mathbb{P}(X = x)$$

Example (Rolling a die)

We can model a roll of a fair die as a random variable X with $\mathbb{P}(X = i) = \frac{1}{6}$ for $i \in \{1, 2, 3, 4, 5, 6\}$. Then $\mathbb{E}[X] =$

Recall that for a set of data $\{x_i\}$, we had $\bar{x} = \frac{\sum f_i x_i}{\sum f_i}$, so we have

Definition. The **variance** of a random variable X is defined to be:

$$\sigma^2 = \text{Var}[X] = \mathbb{E}[(X - \mu)^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$
$$\text{Var}[X] = \sum_x x^2 \mathbb{P}(X = x) - \left(\sum_x x \mathbb{P}(X = x) \right)^2$$

Example (Rolling a die)

We can model a roll of a fair die as a random variable X with $\mathbb{P}(X = i) = \frac{1}{6}$ for $i \in \{1, 2, 3, 4, 5, 6\}$. Then $\text{Var}[X] =$

Fact — For a general function $g : \mathbb{R} \rightarrow \mathbb{R}$,

$$\mathbb{E}[g(X)] = \sum_x g(x) \mathbb{P}(X = x)$$

Coding

Suppose X and Y are random variables, related to each other by $Y = aX + b$, where $a, b \in \mathbb{R}$ are constants. Then

Fact —

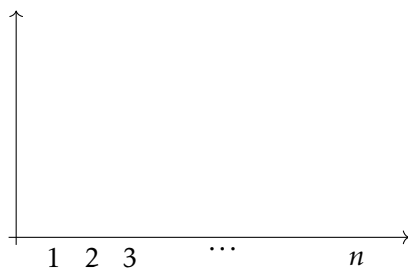
$$\begin{aligned}\mathbb{E}[Y] &= \mathbb{E}[aX + b] = a\mathbb{E}[X] + b \\ \text{Var}[Y] &= \text{Var}[aX + b] = a^2\text{Var}[X]\end{aligned}$$

Discrete Uniform Distribution

Definition. A random variable X is distributed with the **discrete uniform distribution** on $[n] = \{1, 2, \dots, n\}$ if

$$P(X = i) = \begin{cases} \frac{1}{n} & \text{if } i \in \{1, \dots, n\} \\ 0 & \text{otherwise} \end{cases}$$

We often write $X \sim U(n)$, read X is *distributed* with the discrete uniform distribution



Fact — If $X \sim U(n)$

$$\mathbb{E}[X] = \frac{n+1}{2}$$
$$\text{Var}[X] =$$

Binomial Distribution

Example

Suppose we flip a fair coin 5 times, what is the probability we get 3 heads?

Definition. A random variable $X \sim B(n, p)$ is distributed with the **binomial distribution** with n trials and probability p of success if

$$P(X = i) = \begin{cases} \binom{n}{i} p^i (1-p)^{n-i} & \text{if } i \in \{0, 1, \dots, n\} \\ 0 & \text{otherwise} \end{cases}$$

There are two conditions to use a Binomial model:

1. Each trial is either a success or not
2. There is a *fixed* number of trials

We also need to make two assumptions:

1. Each trial has the same probability
2. Each trial is independent.

Example

On how many of the first 10 days of February will it rain?

Example

If I draw 13 cards from a deck of cards, how many red cards will there be?

Example

A fair, five-sided spinner is spun 20 times. What is the probability we get exactly the value 5 exactly twice?

Mean and variance of Binomial distribution

Fact — If $X \sim B(n, p)$, and $q = 1 - p$

$$\begin{aligned}\mu &= \mathbb{E}[X] = np \\ \sigma^2 &= \text{Var}[X] = npq\end{aligned}$$

Linearity of mean and variance**Theorem**

If X and Y are random variables:

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

Theorem

If X and Y are independent random variables (ie $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$):

$$\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$$

Linearity of mean and variance (Attempt 2)

Definition. We say $X \sim \text{Bernoulli}(p)$ or $X \sim B(1, p)$ is a **Bernoulli random variable**

Fact — If $X \sim B(1, p)$, and $q = 1 - p$ then $\mathbb{E}[X] = p, \text{Var}[X] = pq$

Fact — If $X \sim B(n, p)$, and $q = 1 - p$

$$\begin{aligned}\mu &= \mathbb{E}[X] = np \\ \sigma^2 &= \text{Var}[X] = npq\end{aligned}$$

Geometric Distribution

Example

Suppose we are flipping a fair coin until it comes up heads. What is the probability we need k flips?

Definition. A random variable $X \sim \text{Geo}(p)$ is distributed with the **geometric distribution** with probability p of success if

$$P(X = k) = \begin{cases} (1 - p)^{k-1} p & \text{if } k \in \{1, 2, 3, \dots\} \\ 0 & \text{otherwise} \end{cases}$$

Example

A darts player must start a game by hitting a certain section of the dart-board, The probability that any single dart hits this section is $\frac{2}{5}$. The player keeps throwing darts until he hits the required section. Calculate the probability that he needs:

- (a) exactly 5 throws,
- (b) at most 2 throws,
- (c) at least 8 throws,
- (d) at most 10 throws.
- (e) How many throws does the player need to make to be at least 99.999% sure of hitting the required section?
- (f) What assumptions have you made in using the geometric distribution to model this situation?

Fact — If $X \sim \text{Geo}(p)$

$$\mu = \mathbb{E}[X] =$$

$$\sigma^2 = \text{Var}[X] =$$

Example

How many throws will the darts player make in expectation?

Negative Binomial Distribution

Example

Suppose we are flipping a fair coin, how many flips do we need in order to have flipped 2 heads?

Definition. A random variable $X \sim \text{NB}(r, p)$ is distributed with the **negative binomial distribution** which gives the probability of k failures until r successes where the probability of success is p

$$P(X = k) =$$

Example

If $X \sim \text{NB}(r, p)$, calculate $\mathbb{E}(X)$ and $\text{Var}(X)$

Hypergeometric Distribution

Example

Suppose we draw 3 cards from a standard deck of 52. What is the probability we get (exactly) 2 aces?

Definition. A random variable $X \sim \text{Hypergeometric}(N, K, n)$ is distributed with the **hypergeometric distribution** which gives the probability of k successes from n draws, *without replacement*, from a population of size N , with K possible successes

$$P(X = k) =$$

Example

If $X \sim \text{Hypergeometric}(N, K, n)$, calculate $\mathbb{E}(X)$ and $\text{Var}(X)$