Recap

- (Ch 3) Random outcomes and events

Today

- (Ch 4) Random variables

Conditional probability and independence

- The definition of conditional probability is

$$P(E_2|E_1) = \frac{P(E_1 \cap E_2)}{P(E_1)}$$

- If E_1 and E_2 are independent, then $P(E_1 \cap E_2) = P(E_1)P(E_2)$, so

$$P(E_2|E_1) = P(E_2)$$

Random variables

- A random variable is function that maps events to real numbers
- Example: Toss a coin. Let random variable X be
  - 0 if the coin comes up heads
  - 1 if the coin comes up tails

$$\text{Let } X = \begin{cases} 0 & \text{with probability } 0.5 \\ 1 & \text{with probability } 0.5 \end{cases}$$

- Random variables have nothing to do with variables in a program, even though we might use a programming variable to represent a random variable.

Random variables: more examples

- Number of pairs in a hand of 5 cards
  - Let a single outcome be the hand of cards $\binom{52}{5}$
  - Each outcome maps to a number from 0 to 2

- Number of electoral votes that a US presidential candidate will win
  - Let a single outcome be the list of votes or non-votes of all registered voters
  - Each outcome maps to a number from 0 to 538

$(\text{\# candidates} + 1)^{\text{\# registered voters}}$

Random variables and events

- Let X be a random variable

- The set of outcomes $\{A : X(A) = x\}$ is an event with probability $P(X = x)$

- Likewise, the set of outcomes $\{A : X(A) \leq x\}$ is an event with probability $P(X \leq x)$

Random variables and events: dice example

- Roll 2 three-sided dice
- How many outcomes? 9
- Define the following random variables
  - Let X be the value of die 1
  - Let Y be the value of die 2
  - Let sum $S = X + Y$
  - Let difference $D = X - Y$

Random variables and events: dice example

- Calculate the following probabilities
  - $P(X = 1)$ $= 1/3$
  - $P(Y \leq 2)$ $= 2/3$
  - $P(S = 5)$ $= 2/9$
  - $P(D \leq -1)$ $= \frac{3}{9} = \frac{1}{3}$

S

X \ Y	1	2	3
1	2	3	4
2	3	4	5
3	4	5	6

D

X \ Y	1	2	3
1	0	-1	-2
2	1	0	-1
3	2	1	0

Probability distribution

- $P(X = x)$ is called the **probability distribution** of X

- $P(X = x)$ is also denoted as $P(x)$ or $p(x)$

- $P(X = x) \geq 0$ for all values that X can take and is 0 everywhere else

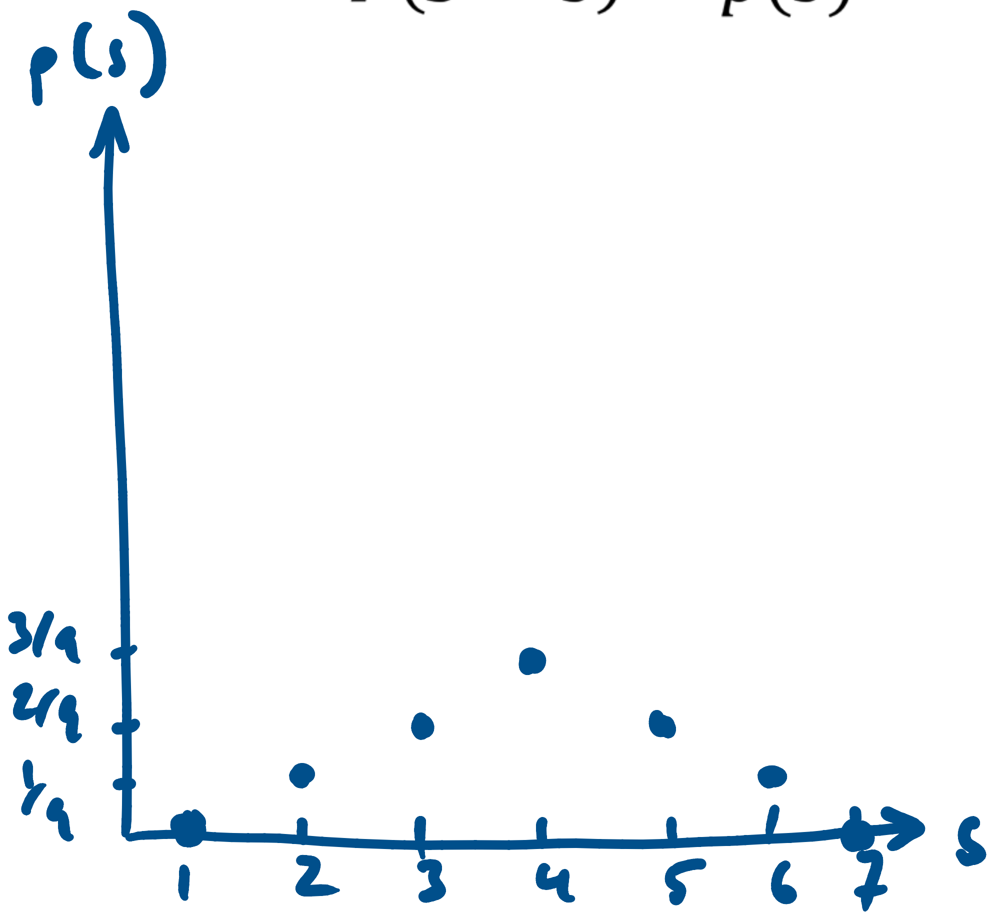
- The sum of the probability distribution $\sum_x P(x) = 1$ because
  - $\{A : X(A) = x_i\}$ and $\{A : X(A) = x_j\}$ are disjoint for $x_i \neq x_j$
  - $\{A : X(A) = x_i\}$ cover the sample space Ω

Cumulative distribution function

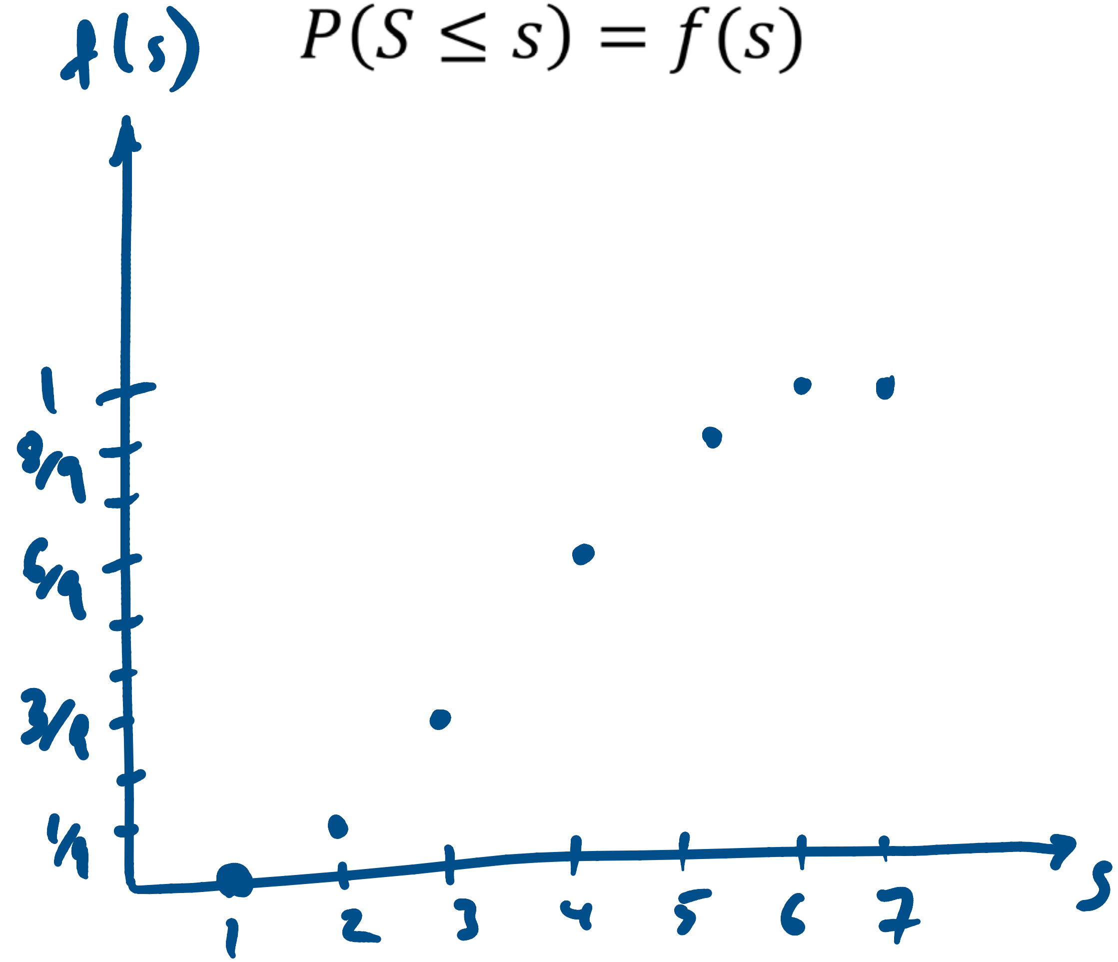
- $P(X \leq x)$ is called the **cumulative distribution function** of X
- $P(X \leq x)$ is also denoted as $f(x)$
- $P(X \leq x)$ is a non-decreasing function of x

Distribution functions: dice example

$$P(S = s) = p(s)$$



$$P(S \leq s) = f(s)$$



Joint probability distribution

- The **joint probability distribution** of two random variables X and Y is
  $P(\{X = x\} \cap \{Y = y\})$, also denoted $P(x, y)$ for short

- We can recover the individual probability distributions $P(x)$ and $P(y)$ from the joint probability distribution as follows
  $P(x) = \sum_y P(x, y)$ and $P(y) = \sum_x P(x, y)$

  marginal distributions

- The sum of the joint probability distribution $\sum_y \sum_x P(x, y) = 1$

Joint probability distribution: dice example

$P(x,y)$

		$y=1$	$y=2$	$y=3$		$P(x)$
x	1	$1/9$	$1/9$	$1/9$	$\rightarrow$	$1/3$
	2	$1/9$	$1/9$	$1/9$	$\rightarrow$	$1/3$
	3	$1/9$	$1/9$	$1/9$	$\rightarrow$	$1/3$
		$\downarrow$	$\downarrow$	$\downarrow$		
$P(y)$		$1/3$	$1/3$	$1/3$		

Joint probability distribution: dice example

$P(s,d)$

s \ d	-2	-1	0	1	2
2	0	0	1/9	0	0
3	0	1/9	0	1/9	0
4	1/9	0	1/9	0	1/9
5	0	1/9	0	1/9	0
6	0	0	1/9	0	0

$P(s)$

1/9
2/9
1/3
2/9
1/9

$P(d)$

1/9	2/9	1/3	2/9	1/9

Independence of random variables

- Random variables X and Y are **independent** if

$$P(x,y) = P(x)P(y) \text{ for all } x \text{ and } y$$

- For the dice example, are the following variables independent?
  - X and Y Yes
  - S and D No e.g. $P(s=2, d=-2) \neq P(s=2)P(d=-2)$

Conditional probability distribution

- The **conditional probability distribution** of X given Y is

$$P(x|y) = \frac{P(x,y)}{P(y)}$$

- For any given y, $\sum_x P(x|y) = 1$

- If X and Y are independent, $P(x,y) = P(x)P(y)$, so $P(x|y) = P(x)$

Conditional distribution: dice example

$$P(s|d) = \frac{P(s,d)}{P(d)}$$

s \ d	-2	-1	0	1	2	p(s)
2	0	0	1/3	0	0	1/9
3	0	1/2	0	1/2	0	2/9
4	1	0	1/3	0	1	1/3
5	0	1/2	0	1/2	0	2/9
6	0	0	1/3	0	0	1/9
p(d)	1/9	2/9	1/3	2/9	1/9	

Bayes rule for random variables

- Bayes rule for events generalizes to random variables

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} = \frac{P(y|x)P(x)}{\sum_x P(y|x)P(x)}$$

- Let's check Bayes rule for a case of the dice example

$$P(D=0|S=2) = \frac{\overset{1/3}{P(S=2|D=0)}\,\overset{1/3}{P(D=0)}}{\underset{1/9}{P(S=2)}} = 1$$