Lecture 12
probability and random variables

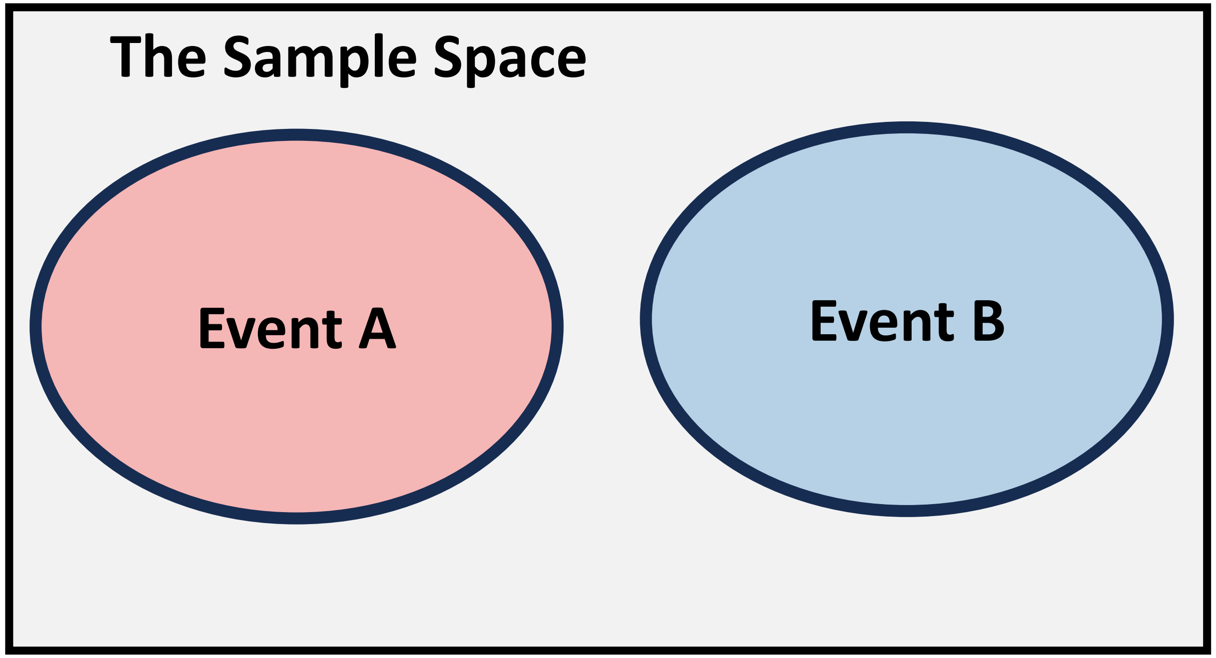
Review

Disjoint vs independent

- Two events A and B are said to be **mutually exclusive/disjoint** when there is no interaction/overlap between them.
  - Mathematically if $A \cap B = \emptyset$
  Then A and B are disjoint

- If two events, A and B are independent, then the outcome of one event has no impact on the outcome of the other event

- Disjoint ≠ independent!!!



Practice:

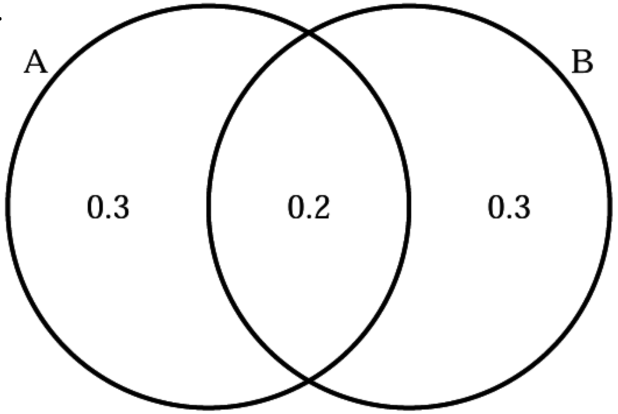
- Ex 1). What is the $P(A \cup B)$?

$$P(A) = 0.2 + 0.3 = 0.5$$
$$P(B) = 0.2 + 0.3 = 0.5$$
$$P(A \cap B) = 0.2$$
$$P(A \cup B) = 0.5 + 0.5 - 0.2 = 0.8$$

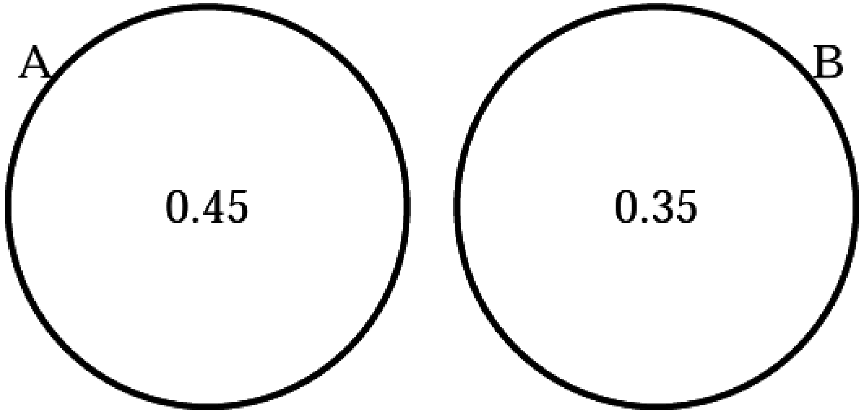
- Ex 2). What is the $P(A \cup B)$?

$$P(A) = 0.45$$
$$P(B) = 0.35$$
$$P(A \cap B) = 0$$
$$P(A \cup B) = 0.45 + 0.35 - 0 = 0.8$$

Ex. 1



Ex. 2



Practice

- Suppose I roll a pair of fair six-sided dice. What is the probability that the roll sums to a value of 8?

$$A = \text{the pair of dice sums to 8}$$

How to find $P(A)$?

Remember

$$P(A) = \frac{\text{\# of ways to get } A}{\text{total \# of possible outcomes}}$$

	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	**(2,6)**
3	(3,1)	(3,2)	(3,3)	(3,4)	**(3,5)**	(3,6)
4	(4,1)	(4,2)	(4,3)	**(4,4)**	(4,5)	(4,6)
5	(5,1)	(5,2)	**(5,3)**	(5,4)	(5,5)	(5,6)
6	(6,1)	**(6,2)**	(6,3)	(6,4)	(6,5)	(6,6)

	2	3	4	5	6	7
(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)	8
(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)	9
(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)	10
(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)	11
(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)	12
(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)	

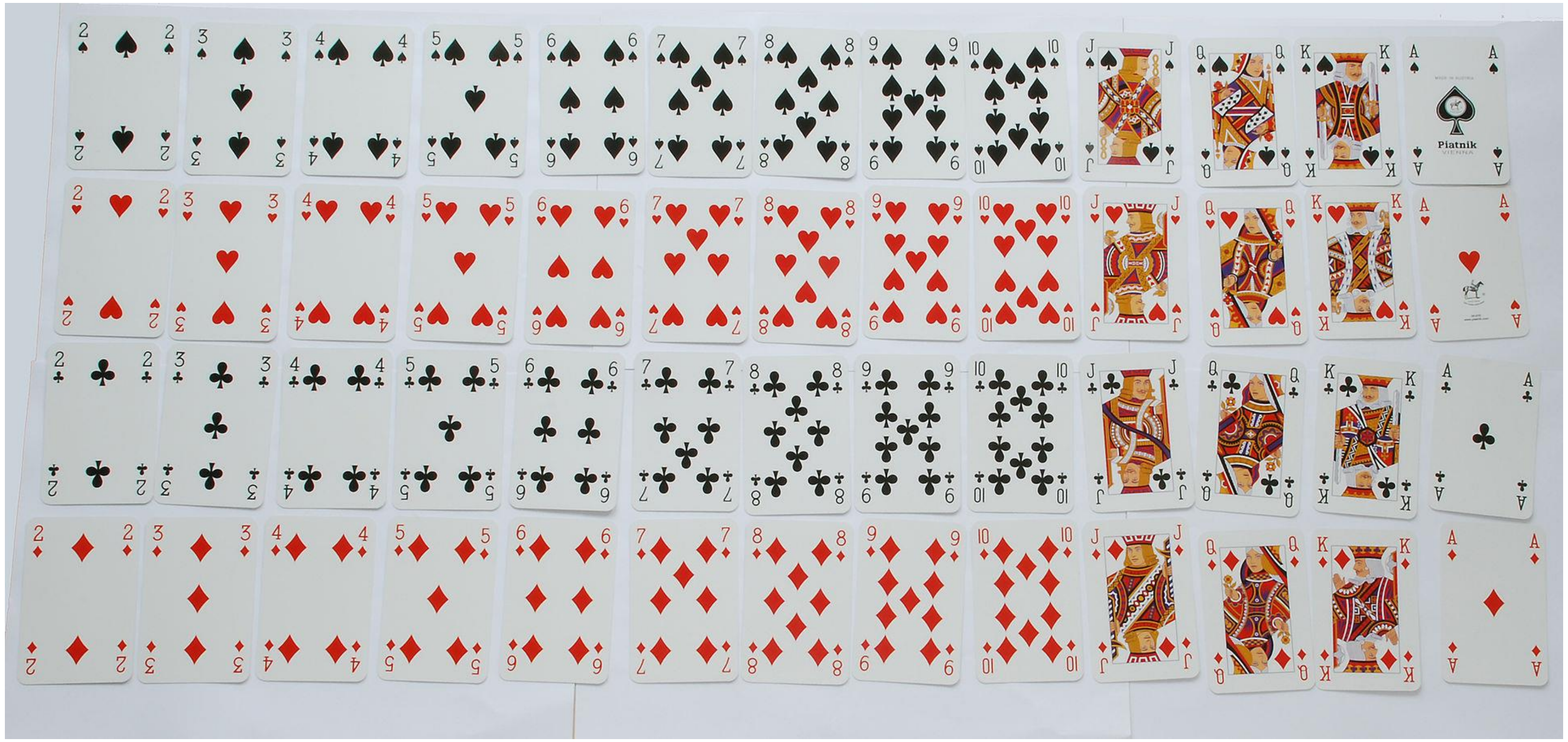
Practice

Suppose a select a card at random from a well-shuffled deck of cards. What is the probability that the card is a king or hearts?

$A =$ card is a king

$B =$ the suit is hearts

Question: What is $P(A \cup B)$



Random Variables and Probability Distributions

- Review: Randomly sampling a population represents a random trial just like rolling a die or flipping a coin.

- The value of a variable measured on a randomly sampled individual is an outcome of a random trial. The following are therefore also random trials:

    Measuring the heights of a randomly selected college students

    Randomly sampling the diameter of trees on a plot of land

- A **Random Variable** occurs when we assign values to the outcomes of random processes. Formally, it is a function that maps from the sample space of event to a value.

    - discrete random variables have a countable number of values

    -continuous random variables have an uncountable number of values

Probability Distributions

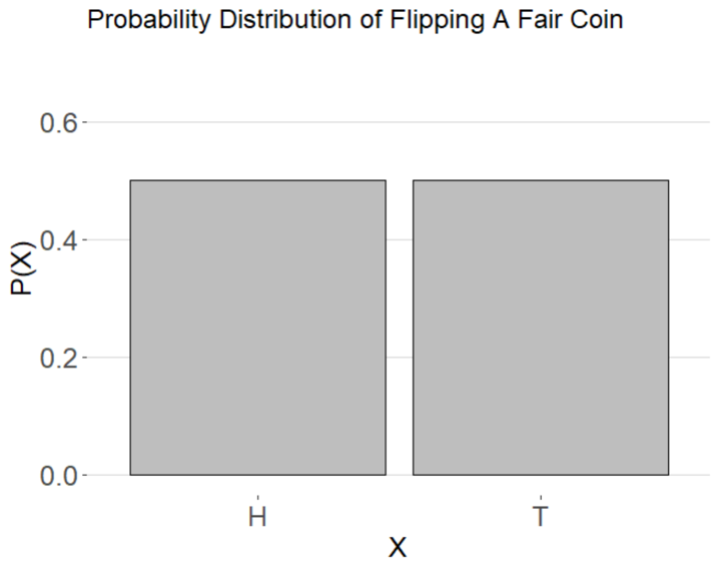
- The distribution of a random variable is called a **probability distribution** – a function that gives the probabilities of different possible outcomes of a random variable.

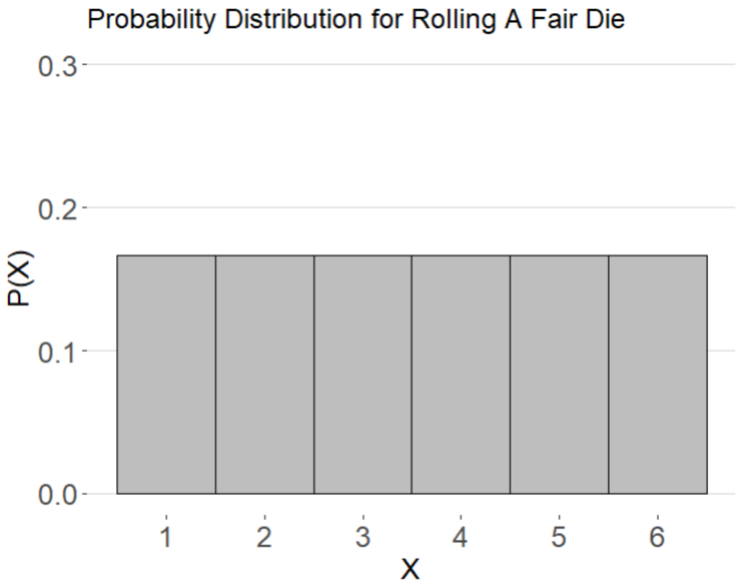
- **Discrete random variables** have a countable number of values such as (0, 1, 2, ….)
  - we will denote a random variable using the capital letters *X* and *Y*

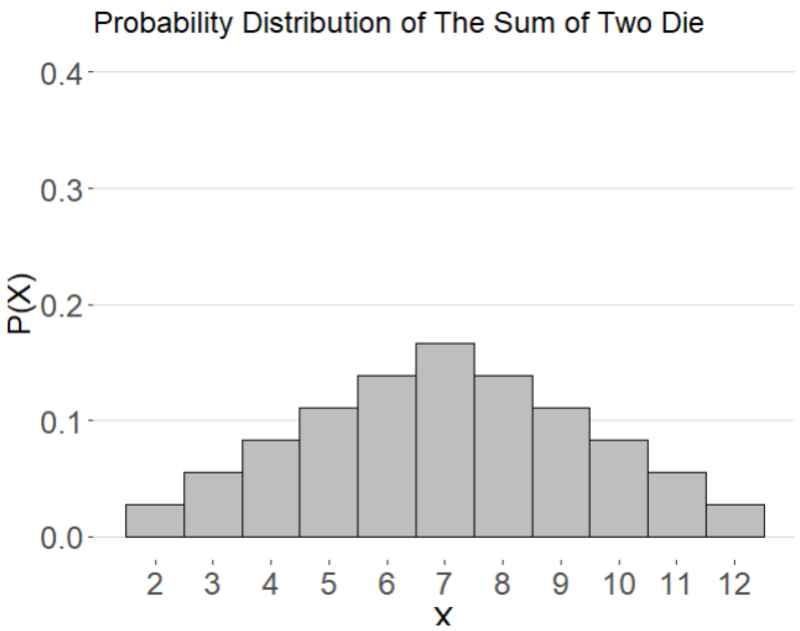
- A random variable is called a **continuous random variable** if the possible values are not countable (more on these later)

- The **probability distribution** of a discrete random variable assigns a probability to each possible outcome
  - for each possible value of $X = \{0, 1, 2 \ldots\}$ the probability $P(X)$ is a value between 0 and 1
  - The sum of the probabilities for all possible values of X equals 1

$$\sum_x P(x) = 1$$







X	$P(X)$
H	0.5
T	0.5

X	P(X)
1	0.167
2	0.167
3	0.167
4	0.167
5	0.167
6	0.167

X	$P(X)$
2	0.03
3	0.06
4	0.08
5	0.11
6	0.14
7	0.17
8	0.14
9	0.11
10	0.08
11	0.06
12	0.03

Computing Probabilities from Discrete Probability Distributions

- Often probabilities concerning a discrete random variables can be computed from its probability distribution using summation.

- That is, if we wish to know the probability of observing a value of the discrete random variable X between a and b we can simply sum the values of X in the given interval

$$P(a \leq X \leq b) = \sum_{a \leq x \leq b} P(x)$$

$$P(X > a) = \sum_{x > a} P(x)$$

Computing Probabilities

- Suppose the following probability distribution gives the probabilities for number of goals scored by the well-known football player Lionel Messi in given match. Assume that goals are independent

X	$P(X)$
0	0.250
1	0.540
2	0.121
3	0.060
4	0.021
5	0.008

- What is the probability that Messi scores More than two goals in a game?

$$P(X > 2) = 0.06 + 0.021 + 0.008 = 0.089$$

- What is the probability that Messi scores at most one goal in a game?

$$P(X \leq 1) = 0.25 + 0.54 = 0.79$$

- What is the probability that Messi scores between 1 and 3 goals in a game?

$$P(1 \leq X \leq 3) = 0.54 + 0.121 + 0.06 = 0.721$$

- What is the probability that Messi scores 2 goals or 5 goals in a game?

$$P(X = 2) \text{ or } P(X = 5) = 0.121 + 0.008 = 0.129$$

Two important probability distributions in statistical inference

A **Population Distribution** – is the probability distribution for a single observation

A **Sampling Distribution** – is the probability distribution of a statistic

Deriving Sampling Distributions

$$S = \begin{bmatrix} T,T,T & = 0,0,0 \\ H,T,T & = 1,0,0 \\ T,H,T & = 0,1,0 \\ T,T,H & = 0,0,1 \\ H,H,T & = 1,1,0 \\ H,T,H & = 1,0,1 \\ T,H,H & = 0,1,1 \\ H,H,H & = 1,1,1 \end{bmatrix}$$

- Imagine flipping a coin three times. We are interested in the number of times the coin comes up Heads
- Suppose the coin is not fair, and $P(\text{Heads}) = 0.4$ and $P(\text{Tails}) = 0.6$
  - We now consider Heads a "success" and Tails a "Failure" – we assign Heads a value of 1 and Tails a value of 0

Deriving Sampling Distributions

Population Distribution

Probability distribution for
a single roll of a 4-sided die

X	$P(X)$
1	0.25
2	0.25
3	0.25
4	0.25

Sampling Distribution

Probability distribution for
the average of $n = 2$ rolls of a 4-sided die

$\bar{x}$	Possible Outcomes	
1	(1,1)	1/16
1.5	(1,2), (2,1)	2/16
2	(1,3), (3,1),(2,2)	3/16
2.5	(3,2), (2,3), (4,1),(1,4)	4/16
3	(3,3), (4,2), (2,4)	3/16
3.5	(3,4),(4,3)	2/16
4	(4,4)	1/16

Mean and Standard Deviation of Discrete Random Variables

- The mean of a probability distribution is defined as

$$\mu = \sum_x xP(x)$$

- The variance and standard deviation of a probability distribution are defined as

$$\sigma^2 = \sum_x (x - \mu)^2 P(x)$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_x (x - \mu)^2 P(x)}$$

Where x denotes an outcome of the random variable X and $P(x)$ denotes the probability of the outcome

The Bernoulli distribution

- The **probability mass function (PMF)** of a discrete random is a function that gives the probability that the variable is exactly equal to some value
- A Bernoulli random variable is on which there are two possible outcomes with probabilities p and $1 - p$
- Whenever we assign the outcomes of a random variable to either "success" or "failure" (1 or 0) we are dealing with a Bernoulli random variable

$$\text{mean} = p$$

$$\text{variance} = p(1-p)$$

$$\text{PMF: } P(X = x) = \begin{cases} p, & \text{if success} \\ (1-p), & \text{else} \end{cases}$$

X	$P(X)$
1 (success)	p
0 (failure)	$1 - p$

The Binomial Distribution

- A discrete distribution which describes the probabilities for the number of successful outcomes in a given number of independent trials where each trial has the same probability of success

  It has two parameters:

  $n =$ the number of trials

  p = the probability of "success" or the probability of the outcome of interest.

  mean = np variance = $np(1-p)$

- It describes the proportion of trials in which a particular outcome of interest occurs

- It is a sum of n independent Bernoulli random variables

- There are many examples of binomial random variables
  - the number of heads observed in n flips of a coin where (each times heads has probability $p = \frac{1}{2}$ of occurring)
  - The proportion of deer with chronic wasting disease (CWD)
  - The number of patients who experience headaches as side of effect of taking a drug

The Binomial Distribution

- Probability Mass Function:

$$P(X = k) = \binom{n}{k} \cdot p^k (1-p)^{n-k}$$

- n is the total number of trials (e.g flips of a coin)
- k successes occur with probability p^k
- $n - k$ failures occur with probability p^{n-k}
- $\binom{n}{k}$ is called the binomial coefficient – it represents the number of ways to arrange k successes in n trials

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

The Poisson Distribution

- A Poisson distribution is a discrete probability distribution. It gives the probability of an event happening a certain number of times k within a given interval of time or space.
- The Poisson distribution has only one parameter, λ (lambda), which is the mean number of events. $\lambda > 0$

Probability Mass Function: $P(X = x) = \frac{\lambda^k e^{-\lambda}}{k!}$

Examples:

The number of traffic accidents at a particular intersection in a given day can be modeled using a Poisson distribution.

The number of defective items produced by a machine in a fixed period of time can be modeled with a Poisson distribution, assuming a constant defect rate.