A Crash Course on Discrete Probability

Events and Probability

Consider a random process (e.g., throw a die, pick a card from a deck)

- Each possible outcome is a **simple event** (or sample point).
- The **sample space** Ω is the set of all possible simple events.
- An **event** is a set of simple events (a subset of the sample space).
- With each simple event E we associate a real number $0 \leq \mathbf{Pr}(E) \leq 1$ which is the **probability** of E.

Probability Space

Definition

A probability space has three components:

1. A sample space Ω, which is the set of all possible outcomes of the random process modeled by the probability space;
2. A family of sets $\mathcal{F}$ representing the allowable events, where each set in $\mathcal{F}$ is a subset of the sample space Ω;
3. A probability function $\mathbf{Pr} : \mathcal{F} \rightarrow \mathbf{R}$, satisfying the definition below.

In a discrete probability space the we use $\mathcal{F} =$ "all the subsets of Ω"

Probability Function

Definition

A probability function is any function $\mathbf{Pr} : \mathcal{F} \to \mathbf{R}$ that satisfies the following conditions:

1. For any event E, $0 \leq \mathbf{Pr}(E) \leq 1$;
2. $\mathbf{Pr}(\Omega) = 1$;
3. For any finite or countably infinite sequence of pairwise mutually disjoint events $E_1, E_2, E_3, \ldots$

$$\mathbf{Pr}\left(\bigcup_{i\geq 1} E_i\right) = \sum_{i\geq 1} \mathbf{Pr}(E_i).$$

The probability of an event is the sum of the probabilities of its simple events.

Examples:

Consider the random process defined by the outcome of rolling a die.

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$$

We assume that all "facets" have equal probability, thus

$$\mathbf{Pr}(1) = \mathbf{Pr}(2) = \ldots.\mathbf{Pr}(6) = 1/6.$$

The probability of the event "odd outcome"

$$= \mathbf{Pr}(\{1, 3, 5\}) = 1/2$$

Assume that we roll two dice:
$\mathcal{S} =$ all ordered pairs $\{(i, j),\ 1 \leq i, j \leq 6\}$.
We assume that each (ordered) combination has probability $1/36$.

Probability of the event "sum $= 2$"

$$\mathbf{Pr}(\{(1,1)\}) = 1/36.$$

Probability of the event "sum $= 3$"

$$\mathbf{Pr}(\{(1,2),(2,1)\}) = 2/36.$$

Let $E_1 =$ "sum bounded by 6",

$$E_1 = \{(1,1),(1,2),(1,3),(1,4),(1,5),(2,1),(2,2),$$

$$(2,3),(2,4),(3,1),(3,2),(3,3),(4,1),(4,2),(5,1)\}$$

$$\mathbf{Pr}(E_1) = 15/36$$

Let $E_2 =$ "both dice have odd numbers", $\mathbf{Pr}(E_2) = 1/4$.

$$\mathbf{Pr}(E_1 \cap E_2) =$$

$$\mathbf{Pr}(\{(1,1),(1,3),(1,5),(3,1),(3,3),(5,1)\}) =$$

$$6/36 = 1/6.$$

Independent Events

Definition

Two events E and F are independent if and only if

$$\mathbf{Pr}(E \cap F) \;=\; \mathbf{Pr}(E) \cdot \mathbf{Pr}(F).$$

Independent Events, examples

Example: You pick a card from a deck.

- $E =$ "Pick an ace"
- $F =$ "Pick a heart"

Example: You roll a die

- $E =$ "number is even"
- $F =$ "number is ≤ 4"

Basically, two events are independent if when one happends it doesn't tell you anything about if the other happened.

Conditional Probability

What is the probability that a random student at University of Ioannina was born in Ioannina.
$E_1 =$ the event "born in Ioannina."
$E_2 =$ the event "a student in UoI."
The conditional probability that a a student at UoI was born in Ioannina is written:

$$\mathbf{Pr}(E_1 \mid E_2).$$

Computing Conditional Probabilities

Definition

The conditional probability that event E occurs given that event F occurs is

$$\mathbf{Pr}(E \mid F) = \frac{\mathbf{Pr}(E \cap F)}{\mathbf{Pr}(F)}.$$

The conditional probability is only well-defined if $\mathbf{Pr}(F) > 0$.

By conditioning on F we restrict the sample space to the set F. Thus we are interested in $\mathbf{Pr}(E \cap F)$ "normalized" by $\mathbf{Pr}(F)$.

Example

What is the probability that in rolling two dice the sum is 8 given that the sum was even?

Example

What is the probability that in rolling two dice the sum is 8 given that the sum was even?

$E_1 =$ "sum is 8",

$E_2 =$ "sum even",

Example

What is the probability that in rolling two dice the sum is 8 given that the sum was even?

$E_1 =$ "sum is 8",

$E_2 =$ "sum even",

$\mathbf{Pr}(E_1) = \mathbf{Pr}(\{(2,6),(3,5),(4,4),(5,3),(6,2)\}) = 5/36$

Example

What is the probability that in rolling two dice the sum is 8 given that the sum was even?

$E_1 =$ "sum is 8",

$E_2 =$ "sum even",

$\mathbf{Pr}(E_1) = \mathbf{Pr}(\{(2,6),(3,5),(4,4),(5,3),(6,2)\}) = 5/36$

$\mathbf{Pr}(E_2) = 1/2 = 18/36.$

Example

What is the probability that in rolling two dice the sum is 8 given that the sum was even?

$E_1 =$ "sum is 8",

$E_2 =$ "sum even",

$$\mathbf{Pr}(E_1) = \mathbf{Pr}(\{(2,6),(3,5),(4,4),(5,3),(6,2)\}) = 5/36$$

$$\mathbf{Pr}(E_2) = 1/2 = 18/36.$$

$$\mathbf{Pr}(E_1 \mid E_2) = \frac{\mathbf{Pr}(E_1 \cap E_2)}{\mathbf{Pr}(E_2)} = \frac{5/36}{1/2} = 5/18.$$

A Useful Identity

Assume two events A and B.

$$\begin{aligned}\Pr(A) &= \Pr(A \cap B) + \Pr(A \cap B^c) \\ &= \Pr(A \mid B) \cdot \Pr(B) + \Pr(A \mid B^c) \cdot \Pr(B^c)\end{aligned}$$

A Useful Identity

Assume two events A and B.

$$\begin{aligned}\Pr(A) &= \Pr(A \cap B) + \Pr(A \cap B^c) \\ &= \Pr(A \mid B) \cdot \Pr(B) + \Pr(A \mid B^c) \cdot \Pr(B^c)\end{aligned}$$

Example:
What is the probability that a random person has height > 1.75?
We choose a random person and let A the event that "the person has height > 1.75."
We want $\Pr(A)$.

A Useful Identity

Assume two events A and B.

$$\begin{aligned} \Pr(A) &= \Pr(A \cap B) + \Pr(A \cap B^c) \\ &= \Pr(A \mid B) \cdot \Pr(B) + \Pr(A \mid B^c) \cdot \Pr(B^c) \end{aligned}$$

Example:
What is the probability that a random person has height > 1.75?
We choose a random person and let A the event that "the person has height > 1.75."
We want $\Pr(A)$.
Assume we know that the probability that a man has height > 1.75 is 54% and that a woman has height > 1.75 is 4%.

A Useful Identity

Assume two events A and B.

$$\begin{aligned}\Pr(A) &= \Pr(A \cap B) + \Pr(A \cap B^c) \\ &= \Pr(A \mid B) \cdot \Pr(B) + \Pr(A \mid B^c) \cdot \Pr(B^c)\end{aligned}$$

Example:
What is the probability that a random person has height > 1.75?
We choose a random person and let A the event that "the person has height > 1.75."
We want $\Pr(A)$.
Assume we know that the probability that a man has height > 1.75 is 54% and that a woman has height > 1.75 is 4%.
Define the event B that "the random person is a man."

Example - a posteriori probability

We are given 2 coins:

- one is a fair coin A
- the other coin, B, has head on both sides

We choose a coin at random, i.e. each coin is chosen with probability $1/2$. We then flip the coin.

Given that we got head, what is the probability that we chose the fair coin A???

Define a sample space of ordered pairs $(coin, outcome)$.
The sample space has three points

$$\{(A,h),(A,t),(B,h)\}$$

$$\mathbf{Pr}((A,h)) = \mathbf{Pr}((A,t)) = 1/4$$

$$\mathbf{Pr}((B,h)) = 1/2$$

Define two events:
$E_1 =$ "Chose coin A".
$E_2 =$ "Outcome is head".

$$\mathbf{Pr}(E_1 \mid E_2) = \frac{\mathbf{Pr}(E_1 \cap E_2)}{\mathbf{Pr}(E_2)} = \frac{1/4}{1/4 + 1/2} = 1/3.$$

Bayes Rule

Another way to compute the same thing:

$$\begin{aligned}
\mathbf{Pr}(E_1 \mid E_2) &= \frac{\mathbf{Pr}(E_1 \cap E_2)}{\mathbf{Pr}(E_2)} \\
&= \frac{\mathbf{Pr}(E_2 \mid E1) \cdot \mathbf{Pr}(E_1)}{\mathbf{Pr}(E_2)} \\
&= \frac{\mathbf{Pr}(E_2 \mid E1) \cdot \mathbf{Pr}(E_1)}{\mathbf{Pr}(E_2 \mid E1)\mathbf{Pr}(E1) + \mathbf{Pr}(E_2 \mid \overline{E1})\mathbf{Pr}(\overline{E1})} \\
&= \frac{1/2 \cdot 1/2}{1/2 \cdot 1/2 + 1 \cdot 1/2} = 1/3.
\end{aligned}$$

Independence

Two events A and B are independent if

$$\mathbf{Pr}(A \cap B) = \mathbf{Pr}(A) \times \mathbf{Pr}(B),$$

or

$$\mathbf{Pr}(A \mid B) = \frac{\mathbf{Pr}(A \cap B)}{\mathbf{Pr}(B)} = \mathbf{Pr}(A).$$

Complement

If event E has probability $\mathbf{Pr}(E)$, then the complement of the event $\overline{E}$ has probability $1 - \mathbf{Pr}(E)$.

Sometimes it is easier to compute this probability. For example:

$E =$ "In 3 rolls of the dice I get at least one 6"

Computing all combinations of events where this is true is complex. What is the complement of E?

Complement

If event E has probability $\mathbf{Pr}(E)$, then the complement of the event $\overline{E}$ has probability $1 - \mathbf{Pr}(E)$.

Sometimes it is easier to compute this probability. For example:

$E =$ "In 3 rolls of the dice I get at least one 6"

Computing all combinations of events where this is true is complex. What is the complement of E?

$\overline{E} =$ "In 3 rolls of the dice I **never** get a 6"

Complement

If event E has probability $\mathbf{Pr}(E)$, then the complement of the event $\overline{E}$ has probability $1 - \mathbf{Pr}(E)$.

Sometimes it is easier to compute this probability. For example:

E = "In 3 rolls of the dice I get at least one 6"

Computing all combinations of events where this is true is complex. What is the complement of E?

$\overline{E}$ = "In 3 rolls of the dice I **never** get a 6"

$$\mathbf{Pr}(\overline{E}) = (1 - 1/6) \cdot (1 - 1/6) \cdot (1 - 1/6)$$

Complement

If event E has probability $\mathbf{Pr}(E)$, then the complement of the event $\overline{E}$ has probability $1 - \mathbf{Pr}(E)$.

Sometimes it is easier to compute this probability. For example:

E = "In 3 rolls of the dice I get at least one 6"

Computing all combinations of events where this is true is complex. What is the complement of E?

$\overline{E}$ = "In 3 rolls of the dice I **never** get a 6"

$$\mathbf{Pr}(\overline{E}) = (1 - 1/6) \cdot (1 - 1/6) \cdot (1 - 1/6)$$

$$\mathbf{Pr}(E) = 1 - \mathbf{Pr}(\overline{E}) = 1 - (5/6)^3$$

Random Variable

Definition

A random variable X on a sample space Ω is a function on Ω; that is, $X : \Omega \to \mathcal{R}$.
A discrete random variable is a random variable that takes on only a finite or countably infinite number of values.

Examples:

In practice, a random variable is some random quantity that we are interested in:

1. I roll a die, $X =$ "result"

Examples:

In practice, a random variable is some random quantity that we are interested in:

1. I roll a die, $X =$ "result"
2. I roll 2 dice, $X =$ "sum of the two values"

Examples:

In practice, a random variable is some random quantity that we are interested in:

1. I roll a die, $X =$ "result"
2. I roll 2 dice, $X =$ "sum of the two values"
3. Consider a gambling game in which a player flips two coins, if he gets heads in both coins he wins \$3, else he losses \$1. The payoff of the game is a random variable.

Examples:

In practice, a random variable is some random quantity that we are interested in:

1. I roll a die, $X =$ "result"
2. I roll 2 dice, $X =$ "sum of the two values"
3. Consider a gambling game in which a player flips two coins, if he gets heads in both coins he wins \$3, else he losses \$1. The payoff of the game is a random variable.
4. I pick a card, $X = \begin{cases} 1, & \text{if card is an Ace} \\ 0, & \text{otherwise} \end{cases}$

Examples:

In practice, a random variable is some random quantity that we are interested in:

1. I roll a die, $X =$ "result"
2. I roll 2 dice, $X =$ "sum of the two values"
3. Consider a gambling game in which a player flips two coins, if he gets heads in both coins he wins \$3, else he losses \$1. The payoff of the game is a random variable.
4. I pick a card, $X = \begin{cases} 1, & \text{if card is an Ace} \\ 0, & \text{otherwise} \end{cases}$
5. I pick 10 random students, $X =$ "average weight"

Examples:

In practice, a random variable is some random quantity that we are interested in:

1. I roll a die, $X =$ "result"
2. I roll 2 dice, $X =$ "sum of the two values"
3. Consider a gambling game in which a player flips two coins, if he gets heads in both coins he wins \$3, else he losses \$1. The payoff of the game is a random variable.
4. I pick a card, $X = \begin{cases} 1, & \text{if card is an Ace} \\ 0, & \text{otherwise} \end{cases}$
5. I pick 10 random students, $X =$ "average weight"
6. $X =$ "Running time of quicksort"

Independent random variables

Definition

Two random variables X and Y are independent if and only if

$$\mathbf{Pr}((X = x) \cap (Y = y)) \quad = \quad \mathbf{Pr}(X = x) \cdot \mathbf{Pr}(Y = y)$$

for all values x and y.

Independent random variables

- A player rolls 5 dice. The sum in the first 3 dice and the sum in the last 2 dice are independent.

Independent random variables

- A player rolls 5 dice. The sum in the first 3 dice and the sum in the last 2 dice are independent.
- I pick a random card from a deck. The value that I got and the suit that I got are independent.

Independent random variables

- A player rolls 5 dice. The sum in the first 3 dice and the sum in the last 2 dice are independent.
- I pick a random card from a deck. The value that I got and the suit that I got are independent.
- I pick a random person in Ioannina. The age and the weight are not independent.

Expectation

Definition

The expectation of a discrete random variable X, denoted by $\mathbf{E}[X]$, is given by

$$\mathbf{E}[X] = \sum_i i \mathbf{Pr}(X = i),$$

where the summation is over all values in the range of X.

Thing of the expectation as the mean value you would get if you took many, many values of the random variable.

Examples:

- The expected value of one die roll is:

$$E[X] = \sum_{i=1}^{6} i\mathbf{Pr}(X = i) = \sum_{i=1}^{6} \frac{i}{6} = 3\frac{1}{2}.$$

- The expectation of the random variable X representing the sum of two dice is

$$\mathbf{E}[X] = \frac{1}{36} \cdot 2 + \frac{2}{36} \cdot 3 + \frac{3}{36} \cdot 4 + \ldots \frac{1}{36} \cdot 12 = 7.$$

- Let X take on the value 2^i with probability $1/2^i$ for $i = 1, 2, \ldots$.

$$\mathbf{E}[X] = \sum_{i=1}^{\infty} \frac{1}{2^i} 2^i = \sum_{i=1}^{\infty} 1 = \infty.$$

Linearity of Expectation

Theorem

For any two random variables X and Y

$$E[X + Y] \ = \ E[X] \ + \ E[Y].$$

Theorem

For any constant c and discrete random variable X,

$$\mathbf{E}[cX] = c\mathbf{E}[X].$$

Note: X and Y do not have to be independent.

Examples:

- The expectation of the sum of n dice is...

Examples:

- The expectation of the sum of n dice is...
- The expectation of the outcome of one die plus twice the outcome of a second die is...

- Assume that N people checked coats in a restaurants. The coats are mixed and each person gets a random coat.
- How many people we expect to have gotten their own coats?

- Assume that N people checked coats in a restaurants. The coats are mixed and each person gets a random coat.
- How many people we expect to have gotten their own coats?
- Let $X =$ "number of people that got their own coats"

- Assume that N people checked coats in a restaurants. The coats are mixed and each person gets a random coat.
- How many people we expect to have gotten their own coats?
- Let $X =$ "number of people that got their own coats"
- It's hard to compute $E[X] = \sum_{k=0}^{N} k\mathbf{Pr}(X = k)$.

- Assume that N people checked coats in a restaurants. The coats are mixed and each person gets a random coat.
- How many people we expect to have gotten their own coats?
- Let X = "number of people that got their own coats"
- It's hard to compute $E[X] = \sum_{k=0}^{N} k\mathbf{Pr}(X = k)$.
- Instead we define N 0-1 random variables X_i:

$$X_i = \begin{cases} 1, & \text{if person } i \text{ got his coat,} \\ 0, & \text{otherwise} \end{cases}$$

- Assume that N people checked coats in a restaurants. The coats are mixed and each person gets a random coat.
- How many people we expect to have gotten their own coats?
- Let X = "number of people that got their own coats"
- It's hard to compute $E[X] = \sum_{k=0}^{N} k\mathbf{Pr}(X = k)$.
- Instead we define N 0-1 random variables X_i:

$$X_i = \begin{cases} 1, & \text{if person } i \text{ got his coat,} \\ 0, & \text{otherwise} \end{cases}$$

- $E[X_i] \ = \ 1 \cdot \mathbf{Pr}(X_i = 1) \ + \ 0 \cdot \mathbf{Pr}(X_i = 0) =$
- $\mathbf{Pr}(X_i = 1) = \dfrac{1}{N}$
- $E[X] \ = \ \displaystyle\sum_{i=1}^{N} E[X_i] \ = \ 1$

Bernoulli Random Variable

A Bernoulli or an indicator random variable:

$$Y = \begin{cases} 1 & \text{if the experiment succeeds,} \\ 0 & \text{otherwise.} \end{cases}$$

$$\mathbf{E}[Y] = p \cdot 1 + (1 - p) \cdot 0 = p = \mathbf{Pr}(Y = 1).$$

Binomial Random Variable

Assume that we repeat n independent Bernoulli trials that have probability p.

Examples:

- I flip n coins, $X_i = 1$, if the ith flip is "head" ($p = 1/2$)
- I roll n dice, $X_i = 1$, if the ith die roll is a 4 ($p = 1/6$)
- I choose n cards, $X_i = 1$, if the ith card is a J, Q, K ($p = 12/52$.)

Let $X = \sum_{i=1}^{n} X_i$.

X is a Binomial random variable.

Binomial Random Variable

Definition

A binomial random variable X with parameters n and p, denoted by $B(n, p)$, is defined by the following probability distribution on $j = 0, 1, 2, \ldots, n$:

$$\mathbf{Pr}(X = j) = \binom{n}{j} p^j (1 - p)^{n-j}.$$

$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!}$ is the number of ways that we can select k elements out of n.

Expectation of a Binomial Random Variable

$$\mathbf{E}[X] = \sum_{j=0}^{n} j \Pr(X = j)$$

$$= \sum_{j=0}^{n} j \binom{n}{j} p^j (1-p)^{n-j}$$

Expectation of a Binomial Random Variable

$$
\begin{aligned}
\mathbf{E}[X] &= \sum_{j=0}^{n} j \Pr(X = j) \\
&= \sum_{j=0}^{n} j \binom{n}{j} p^j (1-p)^{n-j} \\
&= \sum_{j=0}^{n} j \frac{n!}{j!(n-j)!} p^j (1-p)^{n-j} \\
&= \sum_{j=1}^{n} \frac{n!}{(j-1)!(n-j)!} p^j (1-p)^{n-j} \\
&= np \sum_{j=1}^{n} \frac{(n-1)!}{(j-1)!((n-1)-(j-1))!} p^{j-1} (1-p)^{(n-1)-(j-1)} \\
&= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!((n-1)-k)!} p^k (1-p)^{(n-1)-k}
\end{aligned}
$$

Expectation of a Binomial R. V. - 2nd way

Using linearity of expectations

$$\mathbf{E}[X] = \mathbf{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbf{E}[X_i] = np.$$

Computing Expectation

Consider a discrete random variable X that takes values $1, 2, \ldots, k$. Sometimes is it easier to use the following equation to compute the expectation.

$$\mathbf{E}[X] = \sum_{i=1}^{k} \mathbf{Pr}(X \geq i).$$

Proof?