Probability Review

- A **random variable** is a variable that takes on different values rather than having a fixed value. Often, capital letters denote random variables and lowercase letters denote particular values. The **domain** of a random variable is the set of all values the variable can have. The **distribution** of a random variable X, sometimes denoted $p(X)$, specifies the probability that X takes on particular values. If X is **discrete**, $p(X)$ is a **probability mass function (PMF)**, where $p(x)$ is the probability that $X = x$. If X is **continuous**, $p(X)$ is a **probability density function (PDF)**, where $\int_a^b p(x)dx$ is the probability that $a < X < b$.

- We can **sample** from a random variable's distribution to get individual values. The set of all possible outcomes for the sample is the **sample space**. An **event** is a subset of the sample space. For example, an event could be that the random variable X has a value of 22, and we can write its probability as $P(X = 22)$, or $p(22)$ where p is the PMF of X.

- Negation: $P(\text{not } A)$ can be written as $P(\bar{A})$ or $P(\neg A)$ or $P(\sim A)$.
  For any event A, $P(A) + P(\bar{A}) = 1$. Every event either happens or doesn't.

- A **conditional probability** $P(B|A)$ is the probability that B is true *given* that A is true. The **conditional distribution** $p(Y|X)$ is the distribution of random variable Y given the value of random variable X (note that in this case, p takes values of both X and Y as arguments).

- The distribution of a discrete random variable can be represented in a **conditional probability table**: for example, if X and Y are binary random variables, we could have

$P(X=0)$	$P(X=1)$
0.8	0.2

X	$P(Y=0\|X)$	$P(Y=1\|X)$
0	0.9	0.1
1	0.5	0.5

  and we can read off the table to see that, for example, $P(Y = 1|X = 0) = 0.1$.

- Conjunction: $P(A \text{ and } B)$ can be written as $P(A, B)$ or $P(A \cap B)$.
  $P(A, B) = P(A)P(B|A)$. If A and B are independent, $P(A, B) = P(A)P(B)$.
  A modification gives us Bayes' rule: $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$.

- Two random variables X and Y are **independent** if knowing about X does not give you any information about Y - that is, if $p(Y) = p(Y|X)$. (Note that this definition is symmetric: $p(Y) = p(Y|X) \Leftrightarrow p(X) = p(X|Y)$.)

- Two events A and B are **mutually exclusive** if they cannot both happen - that is, if $P(A, B) = 0$. Note that if A and B are mutually exclusive, they cannot be independent (unless both have zero probability).

- Disjunction: $P(A \text{ or } B)$ can be written as $P(A \cup B)$.
$P(A \cup B) = P(A) + P(B) - P(A, B)$. If A and B are mutually exclusive, $P(A \cup B) = P(A) + P(B)$.

- The **joint distribution** $p(X, Y)$ is the distribution of X and Y together. Using the distributions from the previous example, the joint distribution would be

x	y	$P(X = x, Y = y)$
0	0	$P(X = 0)P(Y = 0 \mid X = 0) = 0.8 \times 0.9 = 0.72$
0	1	$P(X = 0)P(Y = 1 \mid X = 0) = 0.8 \times 0.1 = 0.08$
1	0	$P(X = 1)P(Y = 0 \mid X = 1) = 0.2 \times 0.5 = 0.1$
1	1	$P(X = 1)P(Y = 1 \mid X = 1) = 0.2 \times 0.5 = 0.1$

- If we know a joint distribution, we can compute the **marginal distribution** of each of its variables. The marginal distribution is the distribution after "marginalizing out" the other variables. To compute the marginal distribution of Y in the previous example,
$P(Y = 0) = P((Y = 0, X = 0) \cup (Y = 0, X = 1)) = 0.72 + 0.1 = 0.82$
$P(Y = 1) = P((Y = 1, X = 0) \cup (Y = 1, X = 1)) = 0.08 + 0.1 = 0.18.$

Exercises

1. What is the sample space for a fair coin flip? For a sequence of three coin flips? For a sequence of five coin flips in which at least four flips turn out to be heads? Suppose you are told now that the coin is unfair, with the probability of a toss resulting in heads being 0.01. Which answers, if any, change?

2. Suppose you toss two fair, six-sided dice. What is the probability of their sum being exactly 8? You now discover that the result of the first die is 3. Given this new information, what is the probability of the sum of the two dice being exactly 8?

3. A test for a certain rare disease is assumed to be correct 95% of the time: if a person has the disease, the test results are positive with probability 0.95, and if the person does not have the disease, the test results are negative with probability 0.95. A random person drawn from the population has probability 0.001 of having the disease. Given that the person just tested positive, what is the probability of having the disease?

4. A dentist has given us information about the joint probability of two events, T (true if a patient has a toothache) and C (true if a patient has a cavity.) This information is represented in the following table:

	C	$\neg C$
T	0.05	0.05
$\neg T$	0.1	x

   What is the probability of a patient having neither a toothache nor a cavity? What is the probability of having a toothache, and what is the probability of having a cavity? Are these events independent? Given that a patient has a cavity, what is the probability of also having a toothache?

5. Consider two independent rolls of a fair six-sided die, and the following events:

$$\begin{aligned} A &= \{\text{1st roll is 1, 2, or 3}\} \\ B &= \{\text{1st roll is 3, 4, or 5}\} \\ C &= \{\text{the sum of the two rolls is 9}\} \end{aligned}$$

   Compute $P(A \cap B)$, $P(A \cap C)$, $P(B \cap C)$, and $P(A \cap B \cap C)$. Are these three events independent?

6. You are given 1000 coins. Among them, 1 coin has heads on both sides. The other 999 coins are fair coins. You randomly choose a coin and toss it 10 times. Each time, the coin turns up heads. What is the probability that the coin you chose is the unfair one?