# EECS 203: Discrete Mathematics
# Winter 2024
# FoF Discussion 10 Notes

## 1 Conditional Probability

**Probability of an Event (Equally Likely Outcomes):** The probability of an event $E \subseteq S$ is $P(E) = \frac{|E|}{|S|}$ given all elements in $S$ are equally likely.

**Conditional Probability:** The probability of $E_1$ given $E_2$, denoted $P(E_1|E_2)$, is

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

**Independence:** Events $E_1$ and $E_2$ are independent if and only if:

$$P(E_1 \cap E_2) = P(E_1) \cdot P(E_2)$$

Note that this is NOT the same thing as two events being mutually exclusive.

**Mutually Exclusive:** Two events are mutually exclusive if they cannot occur at the same time. Formally, $P(E_1 \cap E_2) = 0$. If two non-empty events are mutually exclusive they are not independent.

**Conditional Probability and Independence:** If $P(E_1) = P(E_1|E_2)$, then $E_1$ and $E_2$ are independent (since $E_2$ doesn't give you any information on $E_1$). More generally, Events $E_1$, $E_2$, and $E_k$ are independent if

$$P(E_1 \cap E_2 \cap ... \cap E_k) = P(E_1) \cdot P(E_2) \cdots P(E_k)$$

## 1.1 Independent vs Mutually Exclusive Events

Discuss with your group the difference between two events being independent and two events being mutually exclusive. Give an example of each.

**Solution:** Mutually exclusive events can be defined as two events cannot occur at same time whereas independent events occur when one event remains unaffected by the occurrence of the other event. An example of a mutually exclusive event is when a coin is a tossed and there are two events that can occur, either it will be a head or a tail. Hence, the events here are mutually exclusive (they cannot co-occur). But if we take two separate coins and flip them, then the occurrence of head or tail on both the coins are independent to each other (one does not affect the other).

## 1.2 Conditional Probability

A bit string of length four is generated at random so that each of the 16 bit strings of length four is equally likely. What is the probability that it contains at least two consecutive 0s, given that its first bit is a 0? (We assume that 0 bits and 1 bits are equally likely.)

**Solution:** Let $E$ be the event that a bit string of length four contains at least two consecutive 0s, and let $F$ be the event that the first bit of a bit string of length four is a 0. The probability that a bit string of length four has at least two consecutive 0s, given that its first bit is a 0, equals

$$P(E|F) = \frac{P(E \cap F)}{P(F)}.$$

Because $E \cap F = 0000, 0001, 0010, 0011, 0100$, we see that $p(E \cap F) = 5/16$. Because there are eight bit strings of length four that start with a 0, we have $p(F) = 8/16 = 1/2$.

Consequently,

$$P(E|F) = \frac{5/16}{1/2} = 5/8.$$

**Bayes' Theorem:** Suppose that $E$ and $F$ are events from a sample space $S$ such that $p(E) \neq 0$ and $p(F) \neq 0$. Then

$$P(F|E) = \frac{p(E|F)p(F)}{P(E)} = \frac{p(E|F)p(F)}{p(E|F)p(F) + p(E|\overline{F})p(\overline{F})} = \frac{p(E \cap F)}{p(E|F)p(F) + p(E|\overline{F})p(\overline{F})}$$

Which version you use depends on what you are given, but they are all equivalent.

## 1.3 Bayes' Theorem $\star$

An electronics company is planning to introduce a new camera phone. The company commissions a marketing report for each new product that predicts either the success or the failure of the product. Of new products introduced by the company, 60% have been successes. Furthermore, 70% of their successful products were predicted to be successes, while 40% of failed products were predicted to be successes. Find the probability that this new camera phone will be successful if its success has been predicted.

**Solution:**

Let $S$ be the event "Product is a success", and let $A$ be the event that "the product is predicted to be successful." We are given $P(S) = 0.6, P(A|S) = 0.7$, and $P(A|\bar{S}) = 0.4$. We want to find $P(S|A)$, the probability that the product is a success given that the it was predicted to be a success.

Applying Bayes' Rule, we have:

$$\begin{aligned} P(S \mid A) =& \frac{P(A \mid S) \cdot P(S)}{P(A)} \\ =& \frac{P(A \mid S) \cdot P(S)}{P(A \mid S) \cdot P(S) + P(A \mid \bar{S}) \cdot P(\bar{S})} \\ =& \frac{(0.7)(0.6)}{(0.7)(0.6) + (0.4)(1 - 0.6)} \\ =& 0.724 \end{aligned}$$

## 1.4 Conditional Probability $\star$

Jakub has created an app that classifies images as either being a Hot Dog or Not a Hot Dog, and he needs your help for some analysis.

Suppose that 4% of the images in a data set are images of hot dogs. Furthermore, suppose that when Jakub's app classifies an image, 97% of the hot dog images are classified correctly (as hot dogs), and 2% of the images that are not hot dogs are classified incorrectly (as hot dogs). What is the probability that:

a) an image classified as a hot dog is really a hot dog?

b) an image classified as a hot dog is **not** a hot dog?

c) an image classified as "not a hot dog" is a hot dog?

d) an image classified as "not a hot dog" is **not** a hot dog?

**Solution:** Let $A$ be the event that a randomly chosen image in the dataset is a hot dog image. Also, let $T$ be the event that a randomly chosen image classification instance comes out positive (the app identifies the image as a hot dog). Based on the problem, $p(A) = 0.04, p(\overline{A}) = 0.96, p(T|A) = 0.97, p(T|\overline{A}) = 0.02$. From these, we can also say that $p(\overline{T}|A) = 0.03$ and $p(\overline{T}|\overline{A}) = 0.98$.

a) This is asking for $p(A|T)$.

$$p(A|T) = \frac{p(T|A)p(A)}{p(T|A)p(A) + p(T|\overline{A})p(\overline{A})} = \frac{(0.97)(0.04)}{(0.97)(0.04) + (0.02)(0.96)} \approx 0.669$$

b) This is asking for $p(\overline{A}|T)$.
Since this is the probability of the complementary event of (a), we can say $p(\overline{A}|T) = 1 - p(A|T) \approx 1 - 0.669 = 0.331$.

c) This is asking for $p(A|\overline{T})$.

$$p(A|\overline{T}) = \frac{p(\overline{T}|A)p(A)}{p(\overline{T}|A)p(A) + p(\overline{T}|\overline{A})p(\overline{A})} = \frac{(0.03)(0.04)}{(0.03)(0.04) + (0.98)(0.96)} \approx 0.001$$

d) This is asking for $p(\overline{A}|\overline{T})$.
Since this is the probability of the complementary event of (c), we can say $p(\overline{A}|\overline{T}) = 1 - p(A|\overline{T}) \approx 1 - 0.001 = 0.999$.

## 1.5 Spam

Suppose that a spam filter is trained on a set of 1000 spam messages and 400 messages that are not spam. Without reading in anything, let the probability that a randomly chosen message is spam be $\frac{1}{2}$. The word "opportunity" appears in 175 spam messages and 20 messages that are not spam. Would an incoming message be rejected as spam if it contains the word "opportunity" and the threshold for rejecting a message is 0.9?

**Solution:** Let $O$ be the event that that a randomly chosen message contains the word "opportunity" in it. Let $S$ be the event that a randomly chosen message is a spam message. Using what is given in the problem, we know that $p(O|S) = \frac{175}{1000}, p(O) = \frac{195}{1400}$, and $p(S) = \frac{1}{2}$ initially without having read in anything. This question is asking for $p(S|O)$.

$$\begin{aligned} p(S|O) &= \frac{p(O|S) \cdot p(S)}{p(O)} \\ &= \frac{\frac{175}{1000} \cdot \frac{1}{2}}{\frac{195}{1400}} \\ &\approx 0.628 \end{aligned}$$

Therefore, since $0.628 < 0.9$, an incoming message would not be rejected as spam if it contains the word "opportunity".

# 2 Random Variables and Expected Value

**Random Variable**: A random variable is a function from the sample space of an experiment to the set of real numbers. That is, a random variable assigns a real number to each possible outcome.

**Indicator Random Variable:** Let $A$ be an event. Then $I_A$, the indicator random variable of $A$, equals 1 if $A$ occurs and equals 0 otherwise. The expectation of the indicator random variable of $A$ equals the probability of $A$, that is, $E(I_A) = p(A)$.

**Expected Value:** The weighted average of values that a random variable can take on, where each possible value is weighted by its respective probability. It can be found using $E(X) = \sum_{s \in S} p(S) \cdot X(S)$ or $E(X) = \sum_{r \in X(S)} p(X = r) \cdot r$

**Linearity of Expectation:** Expected value is linear. This tells us that the expected value of the sum of random variables is equal to the sum of their individual expected values. *Note that random variables do* ***not*** *need to be independent in order for linearity to apply.*

$$E(X_1 + X_2 + ... + X_n) = E(X_1) + E(X_2) + ... + E(X_n)$$

## 2.1 Using Random Variables and Expected Value

Suppose that a coin is flipped three times. Let $X$ be a random variable representing the total number of heads across the three flips.

(a) What are the possible values of $X$ and which outcome(s) are associated with each value of $X$?

(b) Find $E(X)$?

**Solution:**

(a) $X(HHH) = 3$
$X(HHT) = X(HTH) = X(THH) = 2$
$X(TTH) = X(THT) = X(HTT) = 1$
$X(TTT) = 0$

(b) Because the coin is fair and the flips are independent, the probability of each outcome is 1/8. Consequently,

$$
\begin{aligned}
E(X) &= \sum_{r \in range(X)} p(X = r) \cdot r \\
&= p(X = 3) \cdot 3 + p(X = 2) \cdot 2 + p(X = 1) \cdot 1 + p(X = 0) \cdot 0 \\
&= \frac{1}{8} \cdot 3 + \frac{3}{8} \cdot 2 + \frac{3}{8} \cdot 1 + \frac{1}{8} \cdot 0 \\
&= \frac{3}{8} + \frac{6}{8} + \frac{3}{8} \\
&= \frac{3}{2}
\end{aligned}
$$

## 2.2 Expected Value ⋆

Rebecca is creating a new 15-digit passcode using only the digits 1, 2, and 3. She chooses each digit uniformly at random from that set of 3 digits, {1, 2, 3}. What is the expected number of times that the sequence 3113 appears in her passcode? For example, it appears 4 times in the passcode

$$\overline{311}\,\underline{\overline{3}11}\,\underline{\overline{3}}\,\overline{113}\,2\,\underline{3113}.$$

**Solution:**

First, we look at how many possible locations there are for 3113 in the passcode. Since 3113 is 4 characters long and the passcode is 15 characters long, then the possible starting positions for 3113 are positions 1, 2, 3, ... , 12. (If 3113 are the last 4 digits of the passcode, the starting position for that sequence is position 12.)

Let $I_k$ be an indicator random variable relating to starting position $k$, where $I_k = 1$ if there is a 3113 starting at position $k$ and $I_k = 0$ otherwise. Let $p$ be the probability that there is a 3113 starting at position $k$. Then, for any starting position, $p = (1/3)^4$. Then $E(I_k) = 1 \cdot P(I_k = 1) + 0 \cdot P(I_k = 1) = 1 \cdot (1/3)^4 + 0 = (1/3)^4$

Let $X$ be a random variable representing the number of times 3113 appears in the passcode. Using the indicator random variable, defined above, and linearity of expectation, we have:

$$\begin{aligned} E(X) &= E(I_1 + I_2 + \cdots + I_{12}) \\ &= E(I_1) + E(I_2) + \cdots + E(I_{12}) \\ &= \left(\frac{1}{3}\right)^4 + \left(\frac{1}{3}\right)^4 + ... + \left(\frac{1}{3}\right)^4 \\ &= 12\left(\frac{1}{3}\right)^4 \\ &= \frac{4}{27} \end{aligned}$$

# 3 Bernoulli Trials and the Binomial Distribution

**Bernoulli Trials and the Binomial Distribution:** Each performance of an experiment with two possible outcomes is called a Bernoulli trial. In general, a possible outcome of a Bernoulli trial is called a success or a failure. If $p$ is the probability of a success and $q$ is the probability of a failure, it follows that $p + q = 1$.

The **probability** of exactly $k$ successes in $n$ independent Bernoulli trials, with probability of success $p$ and probability of failure $q = 1 - p$, is

$$P(numSuccesses = k) = \binom{n}{k} \cdot p^k \cdot q^{n-k}$$

The **expected number of successes** when $n$ mutually independent Bernoulli trials are performed, where $p$ is the probability of success on each trial, is

$$E(numSuccesses) = np$$

## 3.1 Bernoulli Trials and the Binomial Distribution (I) $\star$

A biased program is generating a 10-bit string. For any single bit, that the probability that the program generates a 0 is 0.9, and each of the 10 bits are generated independently.

(a) What is the probability that the bitstring will contain exactly 8 0's?

(b) What is the expected number of of 1's in bitstring?

**Solution:**

(a) This can be viewed as 10 Bernoulli Trails, where "success" is generating a 0 bit, so $p = 0.9$. The probability of exactly 8 successes in 10 trials is $\binom{10}{8}p^8(1-p)^2 = \binom{10}{8}(0.9)^8(0.1)^2$.

(b) By linearity of expectation, $E(number\ of\ 1s) = 10 \cdot P(single\ bit\ is\ 1) = 10(0.1) = 1$.

## 3.2 Bernoulli Trials and the Binomial Distribution (II)

A group of six people play the game of "odd person out" to determine who will buy refreshments. Each person flips a fair coin. If there is a person whose outcome is not the same as that of any other member of the group, this person has to buy the refreshments.

(a) What is the probability that there is an odd person out after the coins are flipped once?

(b) Suppose the group continues to play the game until there is an odd person out. In each iteration of the game, every person flips their coin and then the group checks for an odd person out. What is the expected number of iterations until there is finally an odd person out?

**Solution:**

(a) There are two ways that there could be a odd person out after the coins are flipped: all but one person flipped heads (i.e., exactly 1 tail), or all but one person flipped tails (i.e., exactly 1 head). So, the probability of an odd person out is:

$$\begin{aligned}
p(odd\ out) &= p(exactly\ 1\ tail) + p(exactly\ 1\ head) \\
&= \binom{6}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^5 + \binom{6}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^5 \\
&= 2 \cdot 6 \cdot \left(\frac{1}{2}\right)^6 \\
&= \frac{12}{64} \\
&= \frac{3}{16}
\end{aligned}$$

(b) This is a "waiting time" experiment, which is represented with a Geometric Random Variable. Let $X$ be a random variable representing the number of iterations until the first time there is an odd person out. Each iteration of the game is a Bernoulli trial with probability of success $p = 3/16$ (calculated in Part (a)). We repeat the Bernoulli trials until the first success, which means $X$ has a geometric distribution, and $E(X) = 1/p = 16/3$.

# 4 Geometric Distribution

**Geometric Distribution:** A random variable X has a geometric distribution with parameter $p$ if $P(X = k) = (1 - p)^{k-1}p$ for $k = 1, 2, 3, ...$ where $p$ is a real number $0 \leq p \leq 1$. $P(X = k)$ is the probability that $k - 1$ failures occur before the first success occurs when performing sequential Bernoulli trials. That is, it represents the probability that a specified number of trials is run before a successful trial occurs.

**Expected Value of a Geometric Distribution:** If the random variable $X$ has the geometric distribution with parameter $p$, then $E(X) = 1/p$.

## 4.1 No Broken Hearts (Geometric Distribution)

Regan has a standard deck of 52 playing cards. Regan shuffles the deck, draws a card, and checks if the card has a suit of hearts. If the card has a suit of hearts she stops. Otherwise, she puts the card back into the deck and continues to repeat this process until she draws a card that has a suit of hearts.

a) What is the probability that Regan performs exactly 5 draws?

b) What is the probability that Regan will draw no more than 3 cards?

c) What is the expected number of times Regan will draw a card?

**Solution:** If we let random variable $X$ denote the number of cards Regan will draw, $X$ follows a geometric distribution with $p = \frac{1}{4}$.

a) $P(X=5) = (1-p)^{5-1}(p) = (1-\frac{1}{4})^4(\frac{1}{4}) = (\frac{3}{4})^4(\frac{1}{4})$
$= \frac{3^4}{4^5}$

b) $P(X \leq 3) = P(X=1) + P(X=2) + P(X=3) = (p) + (1-p)^1(p) + (1-p)^2(p) = \frac{1}{4} + (\frac{3}{4})(\frac{1}{4}) + (\frac{3}{4})^2(\frac{1}{4})$
$= \frac{1}{4} + \frac{3}{4^2} + \frac{3^2}{4^3}$

Alternate Solution: $P(X \leq 3) = 1 - P(X > 3)$.
$P(X > 3)$ is the probability that the first 3 cards are not hearts (so that Regan does not stop drawing during the first 3 cards). So, $P(X > 3) = (\frac{3}{4})^3$. $P(X \leq 3) = 1 - (\frac{3}{4})^3$

c) $E(X) = \frac{1}{p} = \frac{1}{\frac{1}{4}}$
$= 4$