# CMPT 210: Probability and Computation

Lecture 12

Sharan Vaswani

June 17, 2022

# Logistics

Late submission for Assignment 2 is Tuesday, 21 June.

Solutions will be released after the Tuesday class, and no submissions are allowed after that.

Midterm is on Friday, 24 June. It will be 50 minutes with material from Lectures 1 - 12.

Please go through the slides and the relevant sections of (Meyer, Lehman, Leighton) to prepare.

The midterm will be "easy" – if your concepts are clear, you should be able to get full marks.

If you have questions about any of the material, ask them on Piazza. Or come to office hours.

# Recap

**Random variable**: A random "variable" $R$ on a probability space is a total function whose domain is the sample space $\mathcal{S}$. The codomain is denoted by $V$ (usually a subset of the real numbers), meaning that $R : \mathcal{S} \to V$.

Example: Suppose we toss three independent, unbiased coins. In this case, $\mathcal{S} = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$. $C$ is a random variable equal to the number of heads that appear such that $C(HHT) = 2$.

**Indicator Random Variables**: An indicator random variable corresponding to an event $E$ is denoted as $\mathcal{I}_E$ and is defined such that for $\omega \in E$, $\mathcal{I}_E[\omega] = 1$ and for $\omega \notin E$, $\mathcal{I}_E[\omega] = 0$.

Example: When throwing two dice, if $E$ is the event that both throws of the dice result in a prime number, then $\mathcal{I}_E((2,4)) = 0$ and $\mathcal{I}_E((2,3)) = 1$.

In general, a random variable that takes on several values partitions $\mathcal{S}$ into several blocks where each block is a subset of $\mathcal{S}$ and is therefore an event.

Example: When tossing three coins, $\Pr[C = 2] = \Pr[\{HHT, HTH, THH\}] = \frac{3}{8}$.

# Recap

**Probability density function (PDF)**: Let $R$ be a random variable with codomain $V$. The probability density function of $R$ is the function $\text{PDF}_R : V \to [0, 1]$, such that $\text{PDF}_R[x] = \Pr[R = x]$ if $x \in \text{Range(R)}$ and equal to zero if $x \notin \text{Range(R)}$.

$\sum_{x \in V} \text{PDF}_R[x] = \sum_{x \in \text{Range(R)}} \Pr[R = x] = 1.$

Example: When tossing three coins, $\text{PDF}_C[2] = \Pr[C = 2] = \frac{3}{8}$.

**Cumulative distribution function (CDF)**: The cumulative distribution function of $R$ is the function $\text{CDF}_R : \mathbb{R} \to [0, 1]$, such that $\text{CDF}_R[x] = \Pr[R \leq x]$.

Example: When tossing three coins,
$\text{CDF}_C[2.3] = \Pr[C \leq 2.3] = \Pr[C = 0] + \Pr[C = 1] + \Pr[C = 2] = \frac{7}{8}$.

Importantly, neither $\text{PDF}_R$ nor $\text{CDF}_R$ involves the sample space of an experiment.

# Distributions

Many random variables turn out to have the same PDF and CDF. In other words, even though $R$ and $T$ might be different random variables on different probability spaces, it is often the case that $\text{PDF}_R = \text{PDF}_T$. Hence, by studying the properties of such PDFs, we can study different random variables and experiments.

**Distribution** over a random variable can be fully specified using the cumulative distribution function (CDF) (usually denoted by $F$). The corresponding probability density function (PDF) is denoted by $f$.

**Common (Discrete) Distributions** in Computer Science:

- Bernoulli Distribution
- Uniform Distribution
- Binomial Distribution
- Geometric Distribution

# Bernoulli Distribution

We toss a biased coin such that the probability of getting a heads is $p$. Let $R$ be the random variable such that $R = 1$ when the coin comes up heads[1] and $R = 0$ if the coin comes up tails. $R$ follows the Bernoulli distribution.

The Bernoulli distribution has the PDF $f \colon \{0, 1\} \to [0, 1]$ meaning that Bernoulli random variables take values in $\{0, 1\}$. It can be fully specified by the "probability of success" (of an experiment) $p$ (probability of getting a heads in the example). Formally, $\text{PDF}_R$ is given by:

$$f(1) = p \quad ; \quad f(0) = q := 1 - p.$$

In the example, $\Pr[R = 1] = f(1) = p = \Pr[\text{event that we get a heads}]$.

The corresponding $\text{CDF}_R$ for the Bernoulli distribution is given by $F : \mathbb{R} \to [0, 1]$:

$$\begin{aligned} F(x) &= 0 && \text{(for } x < 0\text{)} \\ &= 1 - p && \text{(for } 0 \leq x < 1\text{)} \\ &= 1 && \text{(for } x \geq 1\text{)} \end{aligned}$$

[1] In class, we used the convention in (Meyer, Lehman, Leighton) that $R = 0$ if we get a heads. Since this is less standard, from now on, we will use the convention $R = 1$ when we get a heads.

# Uniform Distribution

We roll a standard die. Let $R$ be the random variable equal to the number that shows up on the die. $R$ follows the uniform distribution.

A random variable $R$ that takes on each possible value in its codomain $V$ with the same probability is said to be uniform. The uniform distribution can be fully specified by $V$ and has PDF $f : V \to [0, 1]$ such that:

$$f(v) = {}^{1}/_{|V|}. \qquad \text{(for all } v \in V\text{)}$$

In the example, $f(1) = f(2) = \ldots = f(6) = \frac{1}{6}$.

For $n$ elements in $V$ arranged in increasing order – $(v_1, v_2, \ldots, v_n)$, the CDF is:

$$\begin{aligned} F(x) &= 0 && \text{(for } x < v_1\text{)} \\ &= {}^{k}/_{n} && \text{(for } v_k \leq x < v_{k+1}\text{)} \\ &= 1 && \text{(for } x \geq v_n\text{)} \end{aligned}$$

Q: If $X$ has a Bernoulli distribution, when is $X$ also uniform? Ans: When $p = 1/2$

Questions?

# Binomial Distribution

We toss $n$ biased coins independently. The probability of getting a heads for each coin is $p$. Let $R$ be the random variable equal to the number of heads in the $n$ coin tosses. $R$ follows the Binomial distribution.

$V = \{0, 1, 2, \ldots, n\}$. Hence $\mathrm{PDF}_R$ is a function $f : \{0, 1, 2, \ldots, n\} \to [0, 1]$.

Let $E_k$ be the event we get $k$ heads in $n$ tosses. Let $A_i$ be the event we get a heads in toss $i$.

$$E_k = (A_1 \cap A_2 \ldots A_k \cap A_{k+1}^c \cap A_{k+2}^c \cap \ldots \cap A_n^c) \cup (A_1^c \cap A_2 \ldots A_k \cap A_{k+1} \cap A_{k+2}^c \cap \ldots \cap A_n^c) \cup \ldots$$

$$\begin{aligned} \Pr[E_k] &= \Pr[(A_1 \cap A_2 \ldots A_k \cap A_{k+1}^c \cap A_{k+2}^c \cap \ldots \cap A_n^c)] + \Pr[A_1^c \cap A_2 \ldots A_k \cap A_{k+1} \cap \ldots \cap] + \ldots \\ &= \Pr[A_1] \Pr[A_2] \Pr[A_k] \Pr[A_{k+1}^c] \Pr[A_{k+2}^c] \ldots \Pr[A_n^c] + \ldots = p^k (1-p)^{n-k} + p^k (1-p)^{n-k} + \ldots \end{aligned}$$

$$\implies \Pr[E_k] = \binom{n}{k} p^k (1-p)^{n-k}$$

**Sanity check**: Since $\mathrm{PDF}_R[k] = \Pr[E_k]$ and $V = \{0, 1, 2, \ldots, n\}$,

$$\sum_{i \in V} \mathrm{PDF}_R[i] = \sum_{i=0}^{n} \Pr[E_i] = \sum_{i=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p + 1 - p)^n = 1. \quad \text{(Binomial Theorem)}$$

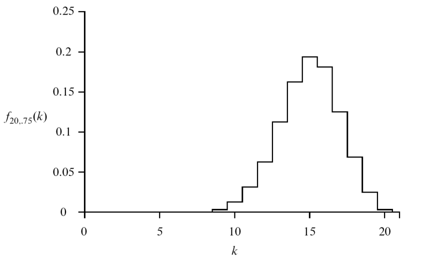
# Binomial Distribution

The binomial distribution can be fully specified by $n, p$ and has PDF $f : \{0, 1, \ldots n\} \to [0, 1]$:

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

The corresponding CDF is given by $F : \mathbb{R} \to [0, 1]$:

$$\begin{aligned} F(x) &= 0 && \text{(for } x < 0\text{)} \\ &= \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} && \text{(for } k \leq x < k+1\text{)} \\ &= 1. && \text{(for } x \geq n\text{)} \end{aligned}$$




Q: If $X$ has a Bernoulli distribution with parameter $p$, does it also follow the Binomial distribution? With what parameters? Ans: Yes. With $n = 1$ and $p = p$.

# Geometric Distribution

We toss a biased coin independently multiple times. The probability of getting a heads is $p$. Let $R$ be the random variable equal to the number of tosses needed to get the first heads. $R$ follows the geometric distribution.

$V = \{1, 2, \ldots, \}$. Hence $\text{PDF}_R$ is a function $f : \{1, 2, \ldots\} \to [0, 1]$.

Let $E_k$ be the event that we need $k$ tosses to get the first heads. Let $A_i$ be the event that we get a heads in toss $i$.

$$\begin{aligned} E_k &= A_1^c \cap A_2^c \cap \ldots \cap A_k \\ \Pr[E_k] &= \Pr[A_1^c \cap A_2^c \cap \ldots \cap A_k] = \Pr[A_1^c] \Pr[A_2^c] \ldots \Pr[A_k] \\ \implies \Pr[E_k] &= (1-p)^{k-1} p \end{aligned}$$

**Sanity check**: Since $\text{PDF}_R[k] = \Pr[E_k]$ and $V = \{1, 2, \ldots, \}$,

$$\sum_{i \in V} \text{PDF}_R[i] = \sum_{i=1}^{\infty} \Pr[E_i] = \sum_{i=1}^{\infty} (1-p)^{i-1} p = \frac{p}{1-(1-p)} = 1. \quad \text{(Sum of geometric series)}$$

# Geometric Distribution

The geometric distribution can be fully specified by $p$ and has PDF $f : \{1, 2, \ldots\} \to [0, 1]$:

$$f(k) = (1 - p)^{k-1} p.$$

The corresponding CDF is given by $F : \mathbb{R} \to [0, 1]$:

$$\begin{aligned} F(x) &= 0 && \text{(for } x < 1\text{)} \\ &= \sum_{i=0}^{k} (1-p)^{i-1} p && \text{(for } k \le x < k+1\text{)} \end{aligned}$$