

CMPT 210: Probability and Computing

Lecture 14

Sharan Vaswani

October 29, 2024

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 $R : \mathcal{S} \rightarrow V$. A r.v partitions the sample space into several blocks.

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 $C : \mathcal{S} \rightarrow \{0, 1, 2, 3\}$. $C(HHT) = 2$.

- For r.v. R , for all $i \in \text{Range}(R)$, the event $[R = i] = \{\omega \in \mathcal{S} | R(\omega) = i\}$. For any r.v. R , $\sum_{i \in \text{Range}(R)} \Pr[R = i] = 1$.

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

$\sum_{i \in \text{Range}(C)} \Pr[C = i] = \Pr[C = 0] + \Pr[C = 1] + \Pr[C = 2] + \Pr[C = 3] = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$.

Recap

- **Indicator Random Variable:** 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 $\mathcal{I}_E((2, 3)) = 1$.

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

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

Importantly, neither PDF_R nor CDF_R involves the sample space of an experiment.

Example: If we flip three coins, and C counts the number of heads, then

$\text{PDF}_C[0] = \Pr[C = 0] = \frac{1}{8}$, and

$\text{CDF}_C[2.3] = \Pr[C \leq 2.3] = \Pr[C = 0] + \Pr[C = 1] + \Pr[C = 2] = \frac{7}{8}$.

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

Canonical Example: 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 and $R = 0$ if the coin comes up tails. R follows the Bernoulli distribution.

PDF_R for Bernoulli distribution: $f: \{0, 1\} \rightarrow [0, 1]$ 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, PDF_R is given by:

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

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

CDF_R for Bernoulli distribution: $F: \mathbb{R} \rightarrow [0, 1]$:

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

Uniform Distribution

Canonical Example: 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.

PDF_R for Uniform distribution: $f : V \rightarrow [0, 1]$ such that for all $v \in V$, $f(v) = 1/|V|$. In the example, $f(1) = f(2) = \dots = f(6) = \frac{1}{6}$.

CDF_R for Uniform distribution: For n elements in V arranged in increasing order – (v_1, v_2, \dots, v_n) , the CDF is:

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

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

Binomial Distribution

Canonical Example: 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.

PDF_R for Binomial distribution: $f : \{0, 1, 2, \dots, n\} \rightarrow [0, 1]$. For $k \in \{0, 1, \dots, n\}$,
 $f(k) = \binom{n}{k} p^k (1-p)^{n-k}$.

Proof: Let E_k be the event we get k heads. Let A_i be the event we get a heads in toss i .

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

$$\Pr[E_k] = \Pr[(A_1 \cap A_2 \dots A_k \cap A_{k+1}^c \cap A_{k+2}^c \cap \dots \cap A_n^c)] + \Pr[A_1^c \cap A_2 \dots A_k \cap A_{k+1} \cap \dots \cap A_n^c] + \dots$$

$$= \Pr[A_1] \Pr[A_2] \Pr[A_k] \Pr[A_{k+1}^c] \Pr[A_{k+2}^c] \dots \Pr[A_n^c] + \dots \quad (\text{Independence of tosses})$$

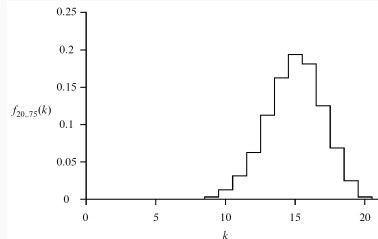
$$= p^k (1-p)^{n-k} + p^k (1-p)^{n-k} + \dots$$

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

(Number of terms = number of ways to choose the k tosses that result in heads = $\binom{n}{k}$)

Binomial Distribution

For the Binomial distribution, $\text{PDF}_R(k) = \binom{n}{k} p^k (1-p)^{n-k}$.



Q: Prove that $\sum_{k \in \text{Range}(R)} \text{PDF}_R[k] = 1$.

By the Binomial Theorem, $\sum_{k \in \text{Range}(R)} \text{PDF}_R[k] = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = (p + 1 - p)^n = 1$.

CDF_R for Binomial distribution: $F : \mathbb{R} \rightarrow [0, 1]$:

$$F(x) = 0 \quad (\text{for } x < 0)$$

$$= \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i} \quad (\text{for } k \leq x < k+1)$$

$$= 1. \quad (\text{for } x \geq n)$$

Geometric Distribution

Canonical Example: 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.

PDF_R for Geometric distribution: $f : \{1, 2, \dots\} \rightarrow [0, 1]$. For $k \in \{1, 2, \dots, \infty\}$,
 $f(k) = (1 - p)^{k-1} p$.

Proof: 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 .

$$E_k = A_1^c \cap A_2^c \cap \dots \cap A_k$$

$$\Pr[E_k] = \Pr[A_1^c \cap A_2^c \cap \dots \cap A_k] = \Pr[A_1^c] \Pr[A_2^c] \dots \Pr[A_k] \quad (\text{Independence of tosses})$$

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

Q: Prove that $\sum_{k \in \text{Range}(R)} \text{PDF}_R[k] = 1$.

By the sum of geometric series, $\sum_{k \in \text{Range}(R)} \text{PDF}_R[k] = \sum_{k=1}^{\infty} (1 - p)^{k-1} p = \frac{p}{1 - (1 - p)} = 1$.

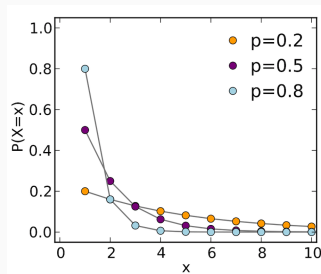
Geometric Distribution

For the Geometric distribution, $\text{PDF}_R(k) = (1 - p)^{k-1}p$.

CDF_R for Geometric distribution: $F : \mathbb{R} \rightarrow [0, 1]$:

$$F(x) = 0 \quad (\text{for } x < 1)$$

$$= \sum_{i=1}^k (1 - p)^{i-1} p \quad (\text{for } k \leq x < k + 1)$$



Questions?

Q: It is known that disks produced by a certain company will be defective with probability 0.01 independently of each other. The company sells the disks in packages of 10 and offers a money-back guarantee that at most 1 of the 10 disks is defective (the package can be returned if there is more than 1 defective disk). What proportion of packages is returned?

Let X be the random variable corresponding to the number of defective disks in a package. Let E be the event that the package is returned. We wish to compute $\Pr[E] = \Pr[X > 1]$. X follows the Binomial distribution $\text{Bin}(10, 0.01)$. Hence,

$$\begin{aligned}\Pr[E] &= \Pr[X > 1] = 1 - \Pr[X \leq 1] = 1 - \Pr[X = 0] - \Pr[X = 1] \\ &= 1 - \binom{10}{0}(0.99)^{10} - \binom{10}{1}(0.99)^9(0.01)^1 \approx 0.05\end{aligned}$$