CMPT 210: Probability and Computing

Lecture 12

Sharan Vaswani

February 15, 2024

Recap - (Basic) Frievald's Algorithm

- **Q**: For $n \times n$ matrices A, B and D, is $D = AB$?
- Last class, we proved that:

	Yes	No
$D = AB$	1	0
$D \neq AB$	$< \frac{1}{2}$	$\geq \frac{1}{2}$

Table 1: Probabilities for Basic Frievalds Algorithm

Frievald's Algorithm

By repeating the *Basic Frievald's Algorithm* m times, we will amplify the probability of success. The resulting complete Frievald's Algorithm is given by:

1. Run the Basic Frievald's Algorithm for m independent runs.
2. If *any* run of the Basic Frievald's Algorithm outputs "no", output "no".
3. If *all* runs of the Basic Frievald's Algorithm output "yes", output "yes".

	Yes	No
$D = AB$	1	0
$D \neq AB$	$< \frac{1}{2^m}$	$\geq 1 - \frac{1}{2^m}$

Table 2: Probabilities for Frievald's Algorithm

If $m = 20$, then Frievald's algorithm will make mistake with probability $1/2^{20} \approx 10^{-6}$.

Computational Complexity: $O(mn^2)$

Probability Amplification

Consider a randomized algorithm $\mathcal{A}$ that is supposed to solve a binary decision problem i.e. it is supposed to answer either Yes or No. It has a one-sided error – (i) if the true answer is Yes, then the algorithm $\mathcal{A}$ correctly outputs Yes with probability 1, but (ii) if the true answer is No, the algorithm $\mathcal{A}$ incorrectly outputs Yes with probability $\leq \frac{1}{2}$.

Let us define a new algorithm $\mathcal{B}$ that runs algorithm $\mathcal{A}$ m times, and if *any* run of $\mathcal{A}$ outputs No, algorithm $\mathcal{B}$ outputs No. If *all* runs of $\mathcal{A}$ output Yes, algorithm $\mathcal{B}$ outputs Yes.

Q: What is the probability that algorithm $\mathcal{B}$ correctly outputs Yes if the true answer is Yes, and correctly outputs No if the true answer is No?

Probability Amplification - Analysis

If A_i denotes run i of Algorithm $\mathcal{A}$, then

$$\Pr[\mathcal{B} \text{ outputs Yes} \mid \text{true answer is Yes}]$$

$$= \Pr[\mathcal{A}_1 \text{ outputs Yes } \cap \mathcal{A}_2 \text{ outputs Yes } \cap \ldots \cap \mathcal{A}_m \text{ outputs Yes} \mid \text{true answer is Yes}]$$

$$= \prod_{i=1}^{m} \Pr[\mathcal{A}_i \text{ outputs Yes} \mid \text{true answer is Yes}] = 1 \qquad \text{(Independence of runs)}$$

$$\Pr[\mathcal{B} \text{ outputs No} \mid \text{true answer is No}]$$

$$= 1 - \Pr[\mathcal{B} \text{ outputs Yes} \mid \text{true answer is No}]$$

$$= 1 - \Pr[\mathcal{A}_1 \text{ outputs Yes } \cap \mathcal{A}_2 \text{ outputs Yes } \cap \ldots \cap \mathcal{A}_m \text{ outputs Yes} \mid \text{true answer is No}]$$

$$= 1 - \prod_{i=1}^{m} \Pr[\mathcal{A}_i \text{ outputs Yes} \mid \text{true answer is No}] \geq 1 - \frac{1}{2^m}.$$

When the true answer is Yes, both $\mathcal{B}$ and $\mathcal{A}$ correctly output Yes. When the true answer is No, $\mathcal{A}$ incorrectly outputs Yes with probability $< \frac{1}{2}$, but $\mathcal{B}$ incorrectly outputs Yes with probability $< \frac{1}{2^m} << \frac{1}{2}$. By repeating the experiment, we have "amplified" the probability of success.

Questions?

Random Variables

Definition: A random "variable" R on a probability space is a total function whose domain is the sample space $\mathcal{S}$. The codomain is usually a subset of the real numbers.

Example: Suppose we toss three independent, unbiased coins. Let C be the number of heads that appear.

$\mathcal{S} = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

C is a total function that maps each outcome in $\mathcal{S}$ to a number as follows: $C(HHH) = 3$, $C(HHT) = C(HTH) = C(THH) = 2$, $C(HTT) = C(THT) = C(TTH) = 1$, $C(TTT) = 0$.

C is a random variable that counts the number of heads in 3 tosses of the coin.

Example: I toss a coin, and define the random variable R which is equal to 1 when I get a heads, and equal to 0 when I get a tails.

Bernoulli random variables: Random variables with the codomain $\{0, 1\}$ are called Bernoulli random variables. E.g. R is a Bernoulli r.v.

Back to throwing dice

Q: Suppose we throw two standard dice one after the other. Let us define R to be the random variable equal to the sum of the dice. What is the domain, range of R?

Ans: $R : \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\} \to \mathbb{N} \cap [2, 12]$.
$R((4, 7)) = 11$, $R((4, 1)) = 5$, $R((1, 1)) = 2$, $R((6, 6)) = 12$.

Q: Three balls are randomly selected from an urn containing 20 balls numbered 1 through 20. The random variable M is the maximal value on the selected balls. What is the domain, range of M? Ans: $M : \{1, 2, \ldots, 20\} \times \{1, 2, \ldots, 20\} \times \{1, 2, \ldots, 20\} \to \{1, 2, \ldots, 20\}$

Q: In the above example, what is $2 \times M((1, 4, 6))$? Is M an invertible function? Ans: 12, No since M maps both $\{1, 2, 5)$ and $(3, 4, 5)$ to 5.

Random Variables and Events

Indicator Random Variable: An indicator random variable maps every outcome to either 0 or 1.

Example: Suppose we throw two standard dice, and define M to be the random variable that is 1 iff both throws of the dice produce a prime number, else it is 0.

$M : \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\} \to \{0, 1\}$. $M((2, 3)) = 1$, $M((3, 6)) = 0$.

An indicator random variable partitions the sample space into those outcomes mapped to 1 and those outcomes mapped to 0.

Example: When throwing two dice, if E is the event that both throws of the dice result in a prime number, then random variable $M = 1$ iff event E happens, else $M = 0$.

The indicator random variable corresponding to an event E is denoted as $\mathcal{I}_E$, meaning that for $\omega \in E$, $\mathcal{I}_E[\omega] = 1$ and for $\omega \notin E$, $\mathcal{I}_E[\omega] = 0$. In the above example, $M = \mathcal{I}_E$ and since $(2, 4) \notin E$, $M((2, 4)) = 0$ and since $(3, 5) \in E$, $M((3, 5)) = 1$.

Random Variables and Events

In general, a random variable that takes on several values partitions $\mathcal{S}$ into several blocks.

Example: When we toss a coin three times, and define C to be the r.v. that counts the number of heads, C partitions $\mathcal{S}$ as follows: $\mathcal{S} = \{\underbrace{HHH}_{C=3}, \underbrace{HHT, HTH, THH}_{C=2}, \underbrace{HTT, THT, TTH}_{C=1}, \underbrace{TTT}_{C=0}\}$.

Each block is a subset of the sample space and is therefore an event. For example, $[C = 2]$ is the event that the number of heads is two and consists of the outcomes $\{HHT, HTH, THH\}$.

Since it is an event, we can compute its probability i.e. $\Pr[C = 2] = \Pr[\{HHT, HTH, THH\}] = \Pr[\{HHT\}] + \Pr[\{HTH\}] + \Pr[\{THH\}]$. Since this is a uniform probability space, $\Pr[\omega] = \frac{1}{8}$ for $\omega \in \mathcal{S}$ and hence $\Pr[C = 2] = \frac{3}{8}$.

Q: What is $\Pr[C = 0]$, $\Pr[C = 1]$ and $\Pr[C = 3]$? Ans: $\frac{1}{8}, \frac{3}{8}, \frac{1}{8}$

Q: What is $\sum_{i=0}^{3} \Pr[C = i]$? Ans: 1

Since a random variable R is a total function that maps every outcome in $\mathcal{S}$ to some value in the codomain, $\sum_{i \in \text{Range of R}} \Pr[R = i] = \sum_{i \in \text{Range of R}} \sum_{\omega \text{ s.t. } R(\omega)=i} \Pr[\omega] = \sum_{\omega \in \mathcal{S}} \Pr[\omega] = 1$.

Back to throwing dice

Q: Suppose we throw two standard dice one after the other. Let us define R to be the random variable equal to the sum of the dice. What are the outcomes in the event $[R = 2]$? Ans: $\{(1, 1)\}$

Q: What is $\Pr[R = 4]$, $\Pr[R = 9]$? Ans: $\frac{3}{36}$, $\frac{4}{36}$

Q: If M is the indicator random variable equal to 1 iff both throws of the dice produces a prime number, what is $\Pr[M = 1]$? Ans: $\frac{9}{36}$

Distribution Functions

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 $x \notin \text{Range(R)}$.

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

Cumulative distribution function (CDF): If the codomain is a subset of the real numbers, then the cumulative distribution function is the function $\text{CDF}_R : \mathbb{R} \to [0,1]$, 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
$\text{CDF}_C[2.3] = \Pr[C \leq 2.3] = \Pr[C = 0] + \Pr[C = 1] + \Pr[C = 2] = \frac{7}{8}$.

Q: What is $\text{CDF}_C[5.8]$? Ans: 1.

For a general random variable R, as $x \to \infty$, $\text{CDF}_R[x] \to 1$ and $x \to -\infty$, $\text{CDF}_R[x] \to 0$.

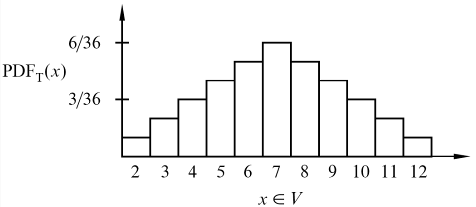
Back to throwing dice

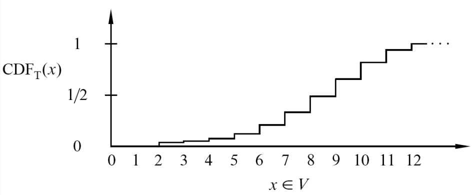
Q: Suppose we throw two standard dice one after the other. Let us define T to be the random variable equal to the sum of the dice. Plot PDF_T and CDF_T

Recall that $T : \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\} \to V$ where $V = \{2, 3, 4, \ldots 12\}$.

$\text{PDF}_T : V \to [0, 1]$ and $\text{CDF}_T : \mathbb{R} \to [0, 1]$.

For example, $\text{PDF}_T[4] = \Pr[T = 4] = \frac{3}{36}$ and $\text{PDF}_T[12] = \Pr[T = 12] = \frac{1}{36}$.





Questions?

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