CMPT 210: Probability and Computation

Lecture 11

Sharan Vaswani

June 14, 2022

Recap

Sample (outcome) space $\mathcal{S}$: Nonempty (countable) set of possible outcomes.

Outcome $\omega \in \mathcal{S}$: Possible "thing" that can happen.

Event E: Any subset of the sample space.

Probability function on a sample space $\mathcal{S}$ is a total function $\Pr : \mathcal{S} \to [0, 1]$. For any $\omega \in \mathcal{S}$,

$$0 \leq Pr[\omega] \leq 1 \quad ; \quad \sum_{\omega \in \mathcal{S}} \Pr[\omega] = 1 \quad ; \quad \Pr[E] = \sum_{\omega \in E} \Pr[\omega]$$

Recap

Union: For mutually exclusive events $E_1, E_2, \ldots, E_n$, $\Pr[E_1 \cup E_2 \cup \ldots E_n] = \Pr[E_1] + \Pr[E_2] + \ldots + \Pr[E_n]$.

Complement rule: $\Pr[E] = 1 - \Pr[E^c]$

Inclusion-Exclusion rule: For any two events E, F, $\Pr[E \cup F] = \Pr[E] + \Pr[F] - \Pr[E \cap F]$.

Union Bound: For any events $E_1, E_2, E_3, \ldots E_n$, $\Pr[E_1 \cup E_2 \cup E_3 \ldots \cup E_n] \leq \sum_{i=1}^{n} \Pr[E_i]$.

Uniform probability space: A probability space is said to be uniform if $\Pr[\omega]$ is the same for every outcome $\omega \in \mathcal{S}$. In this case, $\Pr[E] = \frac{|E|}{|\mathcal{S}|}$.

Recap

Conditional Probability: For events E and F, probability of event E conditioned on F is given by $\Pr[E|F]$ and can be computed as $\Pr[E|F] = \frac{\Pr[E \cap F]}{\Pr[F]}$.

Probability rules with conditioning: For the complement E^c, $\Pr[E^c|F] = 1 - \Pr[E|F]$.

Conditional Probability for multiple events:
$\Pr[E_1 \cap E_2 \cap E_3] = \Pr[E_1] \Pr[E_2|E_1] \Pr[E_3|E_1 \cap E_2]$.

Bayes rule: For events E and F if $\Pr[E] \neq 0$, $\Pr[F|E] = \frac{\Pr[E|F] \Pr[F]}{\Pr[E]}$.

Law of Total Probability: For events E and F, $\Pr[E] = \Pr[E|F] \Pr[F] + \Pr[E|F^c] \Pr[F^c]$.

Independent Events: Events E and F are said to be independent, if knowledge that F has occurred does not change the probability that E occurs, i.e. $\Pr[E|F] = \Pr[E]$ and $\Pr[E \cap F] = \Pr[E] \Pr[F]$.

Independent Events – Generalization to multiple events

Mutual Independence: A set of events is said to be mutually independent if the probability of each event in the set is the same no matter which of the events has occurred.

For any selection of two or more of the events, the probability that all the selected events occur equals the product of the probabilities of the selected events.

Example: For events E_1, E_2 and E_3 to be mutually independent, all the following equalities should hold:

$$\Pr[E_1 \cap E_2] = \Pr[E_1]\Pr[E_2] \quad \Pr[E_1 \cap E_3] = \Pr[E_1]\Pr[E_3]$$
$$\Pr[E_2 \cap E_3] = \Pr[E_2]\Pr[E_3] \quad \Pr[E_1 \cap E_2 \cap E_3] = \Pr[E_1]\Pr[E_2]\Pr[E_3].$$

Can generalize this concept to n events – E_1, E_2, ..., E_n are mutually independent, if for *every* subset of events, the probability that all the selected events occur equals the product of the probabilities of the selected events. Formally, for every subset $S \subseteq \{1, 2, \ldots, n\}$, $\Pr[\cap_{i \in S} E_i] = \prod_{i \in S} \Pr[E_i]$.

Mutual independence vs Pairwise independence

Q: Suppose that we flip three fair, mutually-independent coins. Define the following events: E_1 is the event that coin 1 matches coin 2, E_2 is the event that coin 2 matches coin 3 and E_3 is the event that coin 3 matches coin 1. Are E_1, E_2 and E_3 mutually independent?

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

$\Pr[E_1] = \Pr[\{HHH, HHT, TTH, TTT\}] \implies \Pr[E_1] = \frac{4}{8} = \frac{1}{2}$. Similarly, $\Pr[E_2] = \Pr[E_3] = \frac{1}{2}$.

$\Pr[E_1 \cap E_2] = \Pr[\{HHH, TTT\}] = \frac{2}{8} = \frac{1}{4}$. Hence, $\Pr[E_1 \cap E_2] = \Pr[E_1] \cdot \Pr[E_2]$. Similarly, $\Pr[E_2 \cap E_3] = \Pr[E_2] \cdot \Pr[E_3]$ and $\Pr[E_1 \cap E_3] = \Pr[E_1] \cdot \Pr[E_3]$.

$\Pr[E_1 \cap E_2 \cap E_3] = \Pr[\{HHH, TTT\}] = \frac{2}{8} = \frac{1}{4} \neq \Pr[E_1] \Pr[E_2] \Pr[E_3] = \frac{1}{8}$.

Hence, three events (E_1, E_2 and E_3) can be pairwise independent, but not necessarily mutually independent!

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.

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, codomain of R?

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

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

Example: In the above dice throwing example, let us define M to be the random variable that is 1 iff both throws of the dice produce a prime number, else it is 0.

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

Bernoulli random variables: Random variables with the codomain $\{0, 1\}$ are called Bernoulli random variables. E.g. the indicator random variable.

Random Variables and Events

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 $(3, 5) \in E$, $M((3, 5)) = 1$.

We can define events corresponding to the value that the random variable takes. For example, F is the event that random variable $[R = 3]$, meaning that $F = \{(1, 2), (2, 1)\}$.

If F happens iff $[\mathcal{I}_E = 1]$, then $F = E$.

Random Variables and Events

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

For example, in the coin tossing example, random variable 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: $\frac{9}{36}$

Random Variables - Example

Q: Suppose that an individual purchases two electronic components, each of which may be either defective or acceptable. In addition, suppose that the four possible results — (d, d), (d, a), (a, d), (a, a) — have respective probabilities 0.09, 0.21, 0.21, 0.49 [where (d, d) means that both components are defective, (d, a) that the first component is defective and the second acceptable, and so on]. If we let X be a random variable that denotes the number of acceptable components obtained in the purchase and E be the event that there was at least one acceptable component in the purchase,

- What is the domain, codomain of X? Ans: $\{(d, d), (d, a), (a, d), (a, a)\}$, $\{0, 1, 2\}$
- For every i in the codomain of X, compute $\Pr[X = i]$? Ans: $0.09, 0.42, 0.49$
- What is the domain, codomain of $\mathcal{I}_E$? Ans: $\{(d, d), (d, a), (a, d), (a, a)\}$, $\{0, 1\}$
- For every i in the codomain of $\mathcal{I}_E$, compute $\Pr[\mathcal{I}_E = i]$? Ans: $0.09, 0.91$
- How does X relate to $\mathcal{I}_E$? Ans: $\mathcal{I}_E = \mathcal{I}\{X \geq 1\}$ where $\mathcal{I}$ is the indicator function.

Questions?

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: In the coin tossing example, if 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 $x \to -\infty$, $\text{CDF}_R[x] \to 0$.

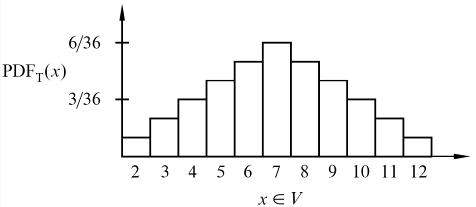
Back to throwing dice

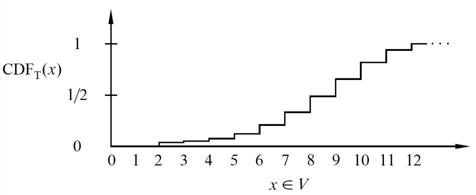
Q: Suppose we throw two standard dice one after the other. Let us define T to be the random variable equal to the sum of the dice. Plot PDF_T and CDF_T

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

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

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





Distribution Functions - Example

Q: Suppose we toss three independent, unbiased coins. Let C be the number of heads that appear. What is PDF_C and CDF_C?

Ans: $\text{PDF}_C[0] = \text{PDF}_C[3] = \frac{1}{8}$, $\text{PDF}_C[1] = \text{PDF}_C[2] = \frac{3}{8}$.
For $x < 0$, $\text{CDF}_C[x] = 0$ and for $x \geq 3$, $\text{CDF}_C[x] = 1$. For $x \in [0,1)$, $\text{CDF}_C[x] = \frac{1}{8}$, for $x \in [1,2)$, $\text{CDF}_C[x] = \frac{4}{8}$ and for $x \in [2,3)$, $\text{CDF}_C[x] = \frac{7}{8}$.

Q: What is $\Pr[1 \leq C \leq 3]$? Ans:
$\Pr[1 \leq C \leq 3] = \Pr[C = 1] + \Pr[C = 2] + \Pr[C = 3] = \text{PDF}_C[1] + \text{PDF}_C[2] + \text{PDF}_C[3] = \frac{7}{8}$.
$\Pr[1 \leq C \leq 3] = \text{CDF}_C[3] - \text{CDF}_C[0] = 1 - \frac{1}{8} = \frac{7}{8}$.

Q: If E is the event that three tosses have the same result, $\text{PDF}_{\mathcal{I}_E}$ and $\text{CDF}_{\mathcal{I}_E}$?

Ans: $\text{PDF}_{\mathcal{I}_E}[1] = 1/4$ and $\text{PDF}_{\mathcal{I}_E}[0] = 3/4$.
For $x < 0$, $\text{CDF}_{\mathcal{I}_E}[x] = 0$, for $x \geq 1$, $\text{CDF}_{\mathcal{I}_E}[x] = 1$ and for $x \in [0,1)$, $\text{CDF}_{\mathcal{I}_E}[x] = 3/4$.