CMPT 210: Probability and Computing

Lecture 13

Sharan Vaswani

February 27, 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 $R : \mathcal{S} \rightarrow V$.

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 $C : \mathcal{S} \rightarrow \{0, 1, 2, 3\}$. $C(HHT) = 2$. An random variable partitions the sample space into several blocks. 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: For the above r.v. C, $[C = 2] = \{HHT, HTH, THH\}$ 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 $\text{PDF}_R : V \to [0,1]$, such that $\text{PDF}_R[x] = \Pr[R = x]$ if $x \in \text{Range(R)}$ and equal to zero if $x \notin \text{Range(R)}$.

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

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

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: $f\colon \{0,1\} \to [0,1]$ 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, 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: $F : \mathbb{R} \to [0,1]$:

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

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: $f : V \to [0,1]$ such that for all $v \in V$, $f(v) = {}^1/_{|V|}$. In the example, $f(1) = f(2) = \ldots = f(6) = \frac{1}{6}$.

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

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

Q: If X has a Bernoulli distribution, when is X also uniform?

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: $f : \{0, 1, 2, \dots, n\} \to [0, 1]$. 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.

$$E_k = (A_1 \cap A_2 \dots A_k \cap A^c_{k+1} \cap A^c_{k+2} \cap \dots \cap A^c_n) \cup (A^c_1 \cap A_2 \dots A_k \cap A_{k+1} \cap A^c_{k+2} \cap \dots \cap A^c_n) \cup \dots$$

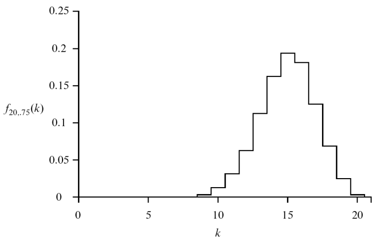
$$\begin{aligned}
\Pr[E_k] &= \Pr[(A_1 \cap A_2 \dots A_k \cap A^c_{k+1} \cap A^c_{k+2} \cap \dots \cap A^c_n)] + \Pr[A^c_1 \cap A_2 \dots A_k \cap A_{k+1} \cap \dots \cap] + \dots \\
&= \Pr[A_1] \Pr[A_2] \Pr[A_k] \Pr[A^c_{k+1}] \Pr[A^c_{k+2}] \dots \Pr[A^c_n] + \dots \qquad \text{(Independence of tosses)} \\
&= p^k (1-p)^{n-k} + p^k (1-p)^{n-k} + \dots
\end{aligned}$$

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

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

Binomial Distribution

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



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

$\mathbf{CDF}_R$ for Binomial distribution: $F : \mathbb{R} \to [0,1]$:

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

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: $f : \{1, 2, \ldots\} \to [0, 1]$. For $k \in \{1, 2, \ldots, \infty\}$, $f(k) = (1-p)^{k-1}\, p$.

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.

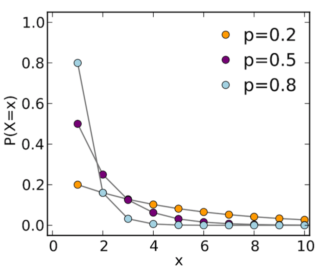
$$\begin{aligned} E_k &= A_1^c \cap A_2^c \cap \ldots \cap A_k \\ \Pr[E_k] &= \Pr[A_1^c \cap A_2^c \cap \ldots \cap A_k] = \Pr[A_1^c]\Pr[A_2^c]\ldots\Pr[A_k] \quad \text{(Independence of tosses)} \\ \implies \Pr[E_k] &= (1-p)^{k-1} p \end{aligned}$$

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: $F : \mathbb{R} \to [0,1]$:

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

Questions?