Arnav Patri

7 Discrete Probability

A **sample space S** is a set of possible outcomes.
An **event** is a subset of the sample space. An **event E** is a subset of the sample space.
If S is a finite and nonempty sample space of equally likely outcomes and $E \subseteq S$, the *probability* of E is given by

$$P(E) = \frac{|E|}{|S|}$$

7.2 Probability Theory

The probability of the **complement** of E, defined as $\bar{E} = S - E$, is found as

$$P(\bar{E}) = 1 - P(E)$$

The **intersection** of 2 events E_1 and E_2, denoted by $E_1 \cap E_2$, is the event that they both occur.
Two events are **disjoint** if the probability of their intersection is 0.
The **union** of 2 events E_1 and E_2, denoted $E_1 \cup E_2$, is the event that exactly one occurs. It is found as

$$P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$$

Probabilities $P(s)$ can be assigned for each $s \in S$ such that $0 \leq P(s) \leq 1$ for all $s \in S$ and $\sum\limits_{s \in S} P(s) = 1$.
The function P from the element $s \in S$ to the probability $P(s)$ is a **probability distribution**.
The probability of an event E can be found as

$$P(E) = \sum_{s \in E} P(s)$$

If $E_1, E_2, \ldots$ is a sequence of pairwise disjoint events, then

$$P\left(\bigcup_i E_i\right) = \sum_i P(E_i)$$

If E_1 and E_2 are events with $P(E_2) > 0$, the **conditional probability** of E_1 given E_2 is denoted and found as

$$P(E_1|E_2) = \frac{P(E_1 \cap E_2)}{P(E_2)}$$

Two events are **independent** if

$$P(E_1 \cap E_2) = P(E_1) \times P(E_2)$$

A set of events is **pairwise independent** if any given 2 events are independent. It is **mutually independent** if every event is independent from every combination of all other events.
A **Bernoulli trial** is a trial with binary outcomes. The probability of success is p.
The **Binomial distribution** gives the probability of X successes after n independent Bernoulli trials with a constant probability of success p.

$$P(x) = \binom{n}{x} p^x (1-p)^x \qquad \text{(binomial)}$$

($P(x)$ is shorthand for $P(X = x)$).
A **random variable** is a function from a sample space to the set of real numbers; that is, a random variable assigns a real number to every possible outcome.
The **distribution** of a random variable X on a sample space S is the set of all pairs $(x, P(x))$ for all $x \in X(S)$.

7.3 Bayes' Theorem

Bayes' theorem states that if E and F are events with nonzero probabilities,

$$P(F|E) = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|\bar{F})P(\bar{F})}$$

7.4 Expected Value and Variance

The **expected value** of a random variable X in sample space S is

$$\mathbb{E}(X) = \mu_X = \sum_{s \in S} P(s)X(s) = \sum_{x \in X} xP(x)$$

The **deviation** of $x \in X$ is $x - \mathbb{E}(X)$.
The expectation operator $\mathbb{E}$ is linear, meaning

that for 2 random variables X and Y and 2 constants a and b,

$$\mathbb{E}(aX + bY) = a\,\mathbb{E}(X) + b\,\mathbb{E}(Y)$$

It should also be noted that

$$\mathbb{E}(aX + b) = a\,\mathbb{E}(X) + b$$

as the expected value of a constant is simply its value.
The **Geometric distribution** gives the probability of getting the first success on the x^{th} independent Bernoulli trial with fixed probability of success p.

$$P(X) = (1-p)^{k-1}p \tag{geometric}$$

Two random variables X and Y are **independent** if

$$P(x \cap y) = P(x) \times P(y)$$

The **variance** of X is

$$\begin{aligned}\mathbb{V}(X) = \sigma_X^2 &= \sum_{s \in S} (X(s) - \mathbb{E}(X))^2 P(s) \\ &= \sum_{x \in X} (x - \mathbb{E}(x))^2 P(x)\end{aligned}$$

Chebyshev's inequality states that for a random variable X and a real number r,

$$P(|x - \mathbb{E}(x)| \geq r) \leq \frac{\mathbb{V}(X)}{r^2}$$

1 Logic and Proofs