Discrete Probability

Patrick CHen

March 26, 2025

Probability theory is a mathematical framework for quantifying the likelihood of the outcome of a random experiment. An experiment is defined as a procedure that generates a random outcome. All possible outcomes of a random experiment is called a sample space. An event is a subset of the sample space.

Axioms of probability theory

Let $\mathcal{S}$ be a sample space. We assign a number $P(E)$ to each event $E \subseteq \mathcal{S}$ that quantifies how likely it is to happen.

- $P(E)$ is the probability that the event E occurs after running a random experiment.
- P is a function $P : \mathcal{P}(\mathcal{S}) \mapsto \mathbb{R}$

The function P must satisfy the following

- $P(E) \in [0, 1]$
- $P(\mathcal{S}) = 1$
- $P(E_1 \cup E_2) = P(E_1) + P(E_2)$ for disjoint sets E_1 and E_2

Probability

Uniform Probability

Let $\mathcal{S}$ be a sample space of equally likely outcomes.

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

Example

There are 20 distinct balls in a box labeled from 1 to 20. If 4 balls are taken out all at once, what is the probability that the balls 1,2,3,4 are chosen.

$$\begin{aligned}
\mathcal{S} &= \{(x, y, z, w) \in \mathbb{Z}^4 \mid 1 \leq x, y, z, w \leq 20, \text{distinct}\} \\
|\mathcal{S}| &= \binom{20}{4} \\
E &= \{(1, 2, 3, 4)\} \\
|E| &= 1 \\
P(E) &= \frac{1}{\binom{20}{4}}
\end{aligned}$$

Example 2

There are 20 distinct balls in a box labeled from 1 to 20. If one ball is taken out and put back in, what is the probability that balls are 1,2,3,4 in that order.

$$\mathcal{S} = \{(x, y, z, w) \in \mathbb{Z}^4 \mid 1 \leq x, y, z, w \leq 20\}$$

$$|\mathcal{S}| = 20^4$$

$$E = \{(1, 2, 3, 4)\}$$

$$|E| = 1$$

$$P(E) = \frac{|E|}{|\mathcal{S}|} = \frac{1}{20^4}$$

Complements and Union

Let $E \subseteq \mathcal{S}$ be an event. $P(\overline{E}) = 1 - P(E)$

$$\begin{aligned}
E \cup \overline{E} &= \mathcal{S} \\
P(E \cup \overline{E}) &= P(\mathcal{S}) \\
P(E) + P(\overline{E}) &= 1 \\
P(\overline{E}) &= 1 - P(E)
\end{aligned}$$

For two non distinct events $E_1, E_2 \subseteq \mathcal{S}$, then $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

$$\begin{aligned}
E_1 \cup E_2 &= (E_1 - E_1 \cap E_2) \cup (E_1 \cap E_2) \cup (E_2 - E_1 \cap E_2) \\
P(E_1 \cup E_2) &= P(E_1 - E_1 \cap E_2) + P(E_1 \cap E_2) + P(E_2 - E_1 \cap E_2) \\
&= P(E_1) - P(E_1 \cap E_2) + P(E_1 \cap E_2) + P(E_2) - P(E_1 \cap E_2) \\
&= P(E_1) + P(E_2) - P(E_1 \cap E_2)
\end{aligned}$$

Non-Uniform Probabilities

Let $\mathcal{S}$ be a set $\{s_1, s_2, \ldots, s_n\}$. The probability of a outcome is the probability of an event that only contains that outcome.

$$P(s) = P(\{s\})$$

From this we can reach a few conclusions

- For any $E \subseteq \mathcal{S}$, then $P(E) = \sum_{s \in E} P(s)$
- $\sum_{s \in \mathcal{S}} P(s) = 1$

Example

What probabilities should we assign to H and T when a fair coin is flipped

$$\mathcal{S} = \{H, T\}, \quad P(\mathcal{S}) = 1, \quad P(H) + P(T) = 1$$

Since its a fair coin, $P(H) = P(T)$. Thus $P(H) = P(T) = \frac{1}{2}$

Example 2

What probabilities should we assign to H and T when H is twice as likely as T

$$P(H) = 2P(T), \quad P(H) + P(T) = 1, \quad 3P(T) = 1$$

Therefore

$$P(T) = \frac{1}{3}, \quad P(H) = \frac{2}{3}$$

Example 3

A die is designed in a way such that 1 is three times more likely to come up than each other number and other outcomes are equally likely. What is the probability that a odd number comes up.

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$$

$$\begin{aligned}
P(1) = 3P(2) = 3P(3) = 3P(4) = 3P(5) = 3P(6) \\
P(1) + P(2) + P(3) + P(4) + P(5) + P(6) = 1 \\
8P(2) = 1 \\
P(1) = \frac{3}{8}, P(2) = P(3) = P(4) = P(5) = P(6) = \frac{1}{8}
\end{aligned}$$

$$\begin{aligned}
P(\text{Odd}) &= P(1) + P(3) + P(5) \\
&= \frac{3}{8} + \frac{1}{8} + \frac{1}{8} \\
&= \frac{5}{8}
\end{aligned}$$

Conditional Probability

Suppose that an event F has already happened. The probability of another event E happening is denoted as $P(E|F)$. Formally, if E and F are two events and $P(F) > 0$, then the probability of E happening given that F happened is defined as follows.

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

Example

Suppose we flip a fair coin three times and we know that the first flip comes up tails. What is the probability that we get an odd number of tails.

$$\begin{aligned} \mathcal{S} &= \{THH, THT, TTH, TTT\} \\ E &= \{THT, TTT\} \\ P(E) &= \frac{2}{4} = \frac{1}{2} \end{aligned}$$

Example 2

What is the probability the sum of two dice rolls is 8 given that the first die shows 3.

$$\begin{aligned} \mathcal{S} &= \{(x, y) \in \mathbb{Z}^2 | 1 \leq x, y \leq 6\} \\ F &= \{(3,1), (3,2), (3,3), (3,4), (3,5), (3,6)\} \\ E &= \{(2,6), (3,5), (4,4), (5,3), (6,2)\} \\ E \cap F &= \{(3,5)\} \\ P(F) &= \frac{6}{36} = \frac{1}{6} \\ P(E \cap F) &= \frac{1}{36} \\ P(E|F) &= \frac{(\frac{1}{36})}{(\frac{1}{6})} = \frac{1}{6} \end{aligned}$$

Example 3

Suppose 36% of families own a dog, 30% of families own a cat, and 22% of families own both. What is the probability that a family owns a dog if they own a cat.

$$\begin{aligned} F &= \text{has cat} = 30\% \\ E &= \text{has dog} = 36\% \\ E \cap F &= \text{has both} = 22\% \\ P(E|F) &= \frac{22\%}{30\%} = \frac{11}{15} \end{aligned}$$

Independence

Two events are independent if $P(E|F) = P(E)$. Equivalently, E and F are independent if and only if $P(E \cap F) = P(E)P(F)$.

- For any arbitrary number of events $E_1, \ldots, E_n$ are pairwise independent if for all pairs of integers i, j such that $i \neq j$,

$$P(E_i \cap E_j) = P(E_i)P(E_j)$$

- For any arbitrary number of events $E_1, \ldots, E_n$ are mutually independent if

$$P(E_1 \cap E_2 \cap \cdots \cap E_m) = P(E_1)P(E_2) \ldots P(E_n)$$

Example

Let E be the event that a bit string of length 4 generated from uniform randomness begins with 1 and F be the event that the bit string contains an even number of 1s. Are E and F independent?

$$
\begin{aligned}
E &= \{1000, 1001, 1010, \ldots, 1111\} = 8 \\
F &= \{0000, 0011, 0101, \ldots, 1111\} = 8 \\
P(E) = P(F) &= \frac{8}{16} = \frac{1}{2} \\
E \cup F &= \{1001, 1010, 1100, 1111\} = 4 \\
P(E \cup F) &= \frac{4}{16} = \frac{1}{4} = (\frac{1}{2})(\frac{1}{2}) = P(E)P(F)
\end{aligned}
$$

Therefore E and F are independent.

Bernoulli Trial

A Bernoulli trial is an experiment with two outcomes: Success with probability p, and failure with probability $1-p$. In n independent trials, the probability of obtaining exactly k success is the product of the amount of choices for how the successes are distributed and the probability that an ordered sequence has k successes.

$$\binom{n}{k} p^k (1-p)^{n-k}$$

Random Variable

A random variable X over a sample space $\mathcal{S}$ is a function $X : \mathcal{S} \mapsto \mathbb{R}$. The distribution of X is the set of pairs $(r, P(X = r))$ where $r \in X(\mathcal{S})$ and $P(X = r)$ is the probability that X has the same value as r.

Example

A coin is biased so that the probability of heads is 2/3. What is the probability that exactly four heads come up when the coin is flipped seven times? Assume coin flips are independent.

$$P(E) = \binom{7}{4} \left(\frac{2}{3}\right)^4 \left(\frac{1}{3}\right)^3$$

Example

Suppose a coin is flipped three times and the random variable X counts the amount of heads. Find the probability distribution of X.

$$
\begin{aligned}
X(HHH) &= 3 \\
X(HHT) = X(HTH) = X(THH) &= 2 \\
X(TTH) = X(THT) = X(HTT) &= 1 \\
X(TTT) &= 0
\end{aligned}
$$

$$
\begin{aligned}
P(X=0) &= \frac{1}{8} \\
P(X=1) &= \frac{3}{8} \\
P(X=2) &= \frac{3}{8} \\
P(X=3) &= \frac{1}{8}
\end{aligned}
$$

The probability distribution of X is

$$
\left\{\left(0, \frac{1}{8}\right), \left(1, \frac{3}{8}\right), \left(2, \frac{3}{8}\right), \left(\frac{3, 1}{8}\right)\right\}
$$

Birthday Problem

How many people need to be in a room so that the probability that at least two of them have the same birthday is greater than $\frac{1}{2}$?

$$
\begin{aligned}
p_1 &= 1 \\
p_2 &= \frac{364}{265} \\
p_2 &= \frac{363}{265} \\
&\vdots \\
p_n &= \frac{365 - k + 1}{365}
\end{aligned}
$$

Let E_n be the event that everyone has a distinct birthday in a room of n people.

$$
P(E_n) = \prod_{k=1}^{n} \frac{365 - k + 1}{365}
$$

$$
\begin{aligned}
P(E_2 2) &\approx 0.524 \\
P(E_2 3) &\approx 0.492
\end{aligned}
$$

Therefore it takes 23 people for the probability of two people having the same birth day to be greater than $\frac{1}{2}$

Bayes Theorem

Let E and F be two events from a sample space $\mathcal{S}$ such that $P(E) \neq 0$ and $P(F) \neq 0$.

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

Proof:

$$\begin{aligned}
P(E|F) &= \frac{P(E \cap F)}{P(F)} \\
P(E|F)P(F) &= P(E \cap F) \\
P(F|E) &= \frac{P(E \cap F)}{P(E)} \\
P(F|E) &= \frac{P(E|F)P(F)}{P(E)}
\end{aligned}$$

Bayes theorem can also be expressed in this following form

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

Generalized Bayes' Theorem

Let E be an event from the sample space $\mathcal{S}$ and $F_1, \ldots, F_n$ be mutually exclusive events such that the union of all is S. $P(E) \neq 0$ and $P(F_i) \neq 0$.

$$P(F_j|E) = \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^{n} P(E|F_i)P(E_i)}$$

Proof:
Since F_i are disjoint

$$E = \bigcup_{i=1}^{n}(F_i \cap E)$$

$$P(F_i \cap E) = P(E|F_i)P(F_i) \Rightarrow P(E) = \sum_{i=1}^{n} P(E|F_i)P(F_i)$$

From Bayes theorem

$$\begin{aligned}
P(F_j|E) &= \frac{P(E|F_j)P(F_j)}{P(E)} \\
&= \frac{P(E|F_j)P(F_j)}{\sum_{i=1}^{n} P(E|F_i)P(F_i)}
\end{aligned}$$

Example

According to the data collected from a soccer team, they have won 25% of all games in the season. They also played with three strikers in 15% of the games. Three strikers where playing in 50% of the games the team won. What is the probability of winning with three strikers?

$$\begin{aligned}
E &= \text{winning game} \\
F &= \text{playing 3 strikers} \\
P(E) &= \frac{25}{100} = \frac{1}{4} \\
P(F) &= \frac{15}{100} \\
P(F|E) &= \frac{50}{100} = \frac{1}{2} \\
P(E|F) &= ?
\end{aligned}$$

$$\begin{aligned}
P(E \cup F) &= P(E)P(F|E) = \frac{1}{4}\frac{1}{2} = \frac{1}{8} \\
P(E|F) &= P(F|E)/(P(F)) = \frac{\frac{1}{8}}{\frac{15}{100}} = \frac{100}{120} = \frac{5}{6}
\end{aligned}$$

Example 2

There are two boxes. The first contains two green balls and seven red balls. The second container four green balls and three red balls. If own box is chosen at random and a red ball is drawn from the boxes, what is the probability that it was drawn from the first box.

$$\begin{aligned}
E &= \text{choose red ball} \\
F &= \text{choosing first box} \\
\overline{F} &= \text{choosing second box} \\
P(E|F) &= \frac{7}{9} \\
P(E|\overline{F}) &= \frac{3}{7} \\
P(F) = P(\overline{F}) &= \frac{1}{2} \\
P(F|E) &= \frac{(\frac{7}{9})(\frac{1}{2})}{(\frac{3}{7})(\frac{1}{2}) + (\frac{3}{7})(\frac{1}{2})} \approx 0.64
\end{aligned}$$

Random Expectations

The expectation of a random variable X on a sample space $\mathcal{S}$ is the average value of X on the sample space. It is defined as follows.

$$\mathbb{E}[X] = \sum_{s \in \mathcal{S}} P(s)X(s) = \sum_{r \in X(\mathcal{S})} P(X = r)r$$

Let $X_1, \ldots, X_n$ be random variables on $\mathcal{S}$ and $a, b \in \mathbb{R}$. The expectation is linear operator on $X_1, \ldots, X_n$

$$\begin{aligned}
\mathbb{E}[X_1 + \cdots + X_n] &= \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n] \\
\mathbb{E}[aX + b] &= a\mathbb{E}[X] + b
\end{aligned}$$

The deviation of X at s defined as $X(s) - \mathbb{E}[X]$

Lemma

The expected value of a random variable that follows the binomial distribution, with success probability p is np
Proof:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{i=1}^{n} P(X=i)\cdot i \\
&= \sum_{i=1}^{n} \binom{n}{i} p^i (1-p)^{n-i}\cdot i \\
&= \sum_{i=1}^{n} n\binom{n-1}{i-1} p^i (1-p)^{n-i}\cdot i \\
&= np\sum_{i=1}^{n} \binom{n-1}{i-1} p^{i-1} (1-p)^{n-i}\cdot i \\
&= np\sum_{i=0}^{n-1} \binom{n-1}{i} p^{i} (1-p)^{n-i+1}\cdot i \\
&= np(p+1-p)^{n-1} \\
&= np
\end{aligned}$$

Example

Let X be the outcome of a fair die roll. What is $\mathbb{E}[X]$?

$$\begin{aligned}
P(x=j) &= \frac{1}{6} \\
\mathbb{E}[X] &= \frac{1}{6}\cdot 1+\cdots+\frac{1}{6}\cdot 6 = \frac{21}{6} = \frac{7}{2}
\end{aligned}$$

Example 2

What is the expected value of the sum of the numbers that appear when a pair of fair dice is rolled.

$$\begin{aligned}
X_1 &= \text{outcome of first die} \\
X_2 &= \text{outcome of second die} \\
\mathbb{E}[X_1+X_2] &= \mathbb{E}[X_1]+\mathbb{E}[X_2] && \text{linearity} \\
&= 2\mathbb{E}[X_1] && \text{identical dice} \\
&= 2(\frac{7}{3}) && \text{from previous example} \\
&= \frac{14}{3}
\end{aligned}$$

Example 3

What is the expected value of a random variable that follows the binomial distribution with success probability p?

$$\begin{aligned}
\mathbb{E}[X_j] &= P(X = \text{success}) \cdot 1 + P(X = \text{fail}) \cdot 0 = p \\
\mathbb{E}[X_1 + \cdots + X_n] &= \mathbb{E}[X_1] + \cdots + \mathbb{E}[X_n] \\
&= \underbrace{p + \cdots + p}_{n} \\
&= np
\end{aligned}$$

Geometric Distribution

A random variable X on a sample space $\mathcal{S}$ has a geometric distribution with parameter p if

$$P(X = k) = (1-p)^{k-1}p, \quad k \geq 1$$

A random variable with a geometric distribution will have an expected value $\mathbb{E}[X] = \frac{1}{p}$.

Example

Suppose that the probability that a coin comes up heads is p. The coin is flipped repeatedly until it comes up heads. What is the expected number of flips until this coin comes up heads?

$$\begin{aligned}
\mathcal{S} &= \{H, TH, TTH, TTTH, TTTTTH, \dots\} \\
X : S \mapsto \mathbb{R} &= \text{The number of tails flipped} \\
\mathbb{E}[X] &= \sum_{i=0}^{\infty} P(X = i) \cdot i \\
&= \sum_{i=0}^{\infty} p(1-p)^{i-1} \cdot i \\
&= p \sum_{i=0}^{\infty} (1-p)^{i-1} \cdot i \\
&= p \frac{1}{p^2} \\
&= \frac{1}{p}
\end{aligned}$$

Independence of Random Variables

Two random variables X, Y are independent if $P(X = r_1, Y = r_2) = P(X = r_1)P(Y = r_1)$.

Product of Independent Expectations

If X and Y are independent random variables over a sample space S, then $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$.

Proof:

$$\begin{aligned}
\mathbb{E}[XY] &= \sum_{r \in XY(s)} P(XY = r)r \\
&= \sum_{r_1 \in X(S), r_2 \in Y(S)} P(X = r_1, Y = r_2) r_1 r_2 \\
&= \sum_{r_1 \in X(S)} \sum_{r_2 \in Y(S)} P(X = r_1) P(Y = r_2) r_1 r_2 \\
&= \sum_{r_1 \in X(S)} P(X = r_1) r_1 \sum_{r_2 \in Y(S)} P(Y = r_2) r_2 \\
&= \mathbb{E}[X]\mathbb{E}[Y]
\end{aligned}$$

Example

Let X_1 and X_2 be the outcome of two die rolls. Define $X = X_1 + X_2$. are X_1 and X independent?

$$\begin{aligned}
P(X_1 = 1, X = 10) &= 0 \\
P(X_1 = 1) &= \frac{1}{6} \\
P(X = 10) = \frac{3}{36} &= \frac{1}{12} \\
0 &\neq \left(\frac{1}{6}\right)\left(\frac{1}{12}\right)
\end{aligned}$$

Therefore, X and X_1 are not independent.

Example 2

Let X and Y be random variables that count the number of heads and the number of tails when a coin is flipped twice. Are X and Y independent?

$$\begin{gathered}
X(HH) = 2, \quad X(HT) = X(TH) = 1, \quad X(TT) = 0 \\
Y(HH) = 0, \quad Y(HT) = Y(TH) = 1, \quad Y(TT) = 2 \\
XY(HH) = 0, \quad XY(HT) = XY(TH) = 1, \quad XY(TT) = 0
\end{gathered}$$

$$\begin{aligned}
\mathbb{E}[XY] &= \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 0 \\
&= \frac{1}{2} \\
\mathbb{E}[X] = \mathbb{E}[Y] &= \frac{1}{4} \cdot 2 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 2 \\
&= \frac{3}{2} \\
\mathbb{E}[XY] &\neq \mathbb{E}[X]\mathbb{E}[Y]
\end{aligned}$$

Thus, X and Y are not independent.

Variance

The variance $V(X)$ of random variable X is defined as

$$V(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

The standard deviation σ is defined as the square root of the variance.

$$\sigma(X) = \sqrt{V(X)}$$

If X and Y are independent, then variance is linear operator.

Lemma

let $\mu = \mathbb{E}[X]$ then $V(X) = \mathbb{E}[(X-\mu)^2]$

$$\begin{aligned}
\mathbb{E}[(X-\mu)^2] &= \mathbb{E}[X^2 - 2\mu X + \mu^2] \\
&= \mathbb{E}[X^2] - 2\mu\mathbb{E}[X] + \mu^2 \\
&= \mathbb{E}[X^2] - 2\mathbb{E}[X]^2 + \mathbb{E}[X]^2 \\
&= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= V(X)
\end{aligned}$$

Example

What is the variance of a Bernoulli random variable with success probability p?

$$X = \begin{cases} 1, & \text{if } p \\ 0 \end{cases}$$

$$X^2 = \begin{cases} 1^2, & \text{if } p \\ 0^2 \end{cases}$$

$$\mathbb{E}[X] = p, \quad \mathbb{E}[X^2] = p$$

$$\begin{aligned}
V(X) &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 \\
&= p - p^2 & &= p(1-p)
\end{aligned}$$

Example

What is the variance of the value of a fair dice.

$$\begin{aligned}
\mathbb{E}[X] &= \frac{7}{2} \\
P(x^2 = j^2) = P(x = j) &= \frac{1}{6} \\
\mathbb{E}[X^2] &= \frac{1}{6}1^2 + \frac{1}{6}2^2 + \cdots + \frac{1}{6}6^2 \\
&= \frac{91}{6} \\
V(X) &= \left(\frac{91}{6}\right) - \left(\frac{7}{2}\right)^2 \approx 2.9
\end{aligned}$$

Chebyshev's Inequality

Let X be a random variable on a sample space $\mathcal{S}$ with $\mu = \mathbb{E}[X]$ and $r \in \mathbb{R}$.

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

Proof:
Define the event $A = \{s \in S : |X(s) - \mu| \geq r\}$.

$$\begin{aligned}
V(X) &= \mathbb{E}[(X-\mu)^2] \\
&= \sum_{s\in A}(X-\mu)^2 P(s) + \sum_{s\notin A}(X-\mu)^2 P(s) \\
&\geq \sum_{s\in A}(X-\mu)^2 P(s) \\
&\geq \sum_{s\in A} r^2 P(s) = r^2 P(A)
\end{aligned}$$