Probability
The chance something happens
We use models to represent a system
Types of models:
- Deterministic: Outcome is determined by input conditions
- Probabilistic/Stochastic: Outcome can vary; degree of randomness

Building a model:
1. Define the random experiment inherent in the problem,
2. Specify the set of all possible outcomes and events of interest,
3. Specify a probability assignment for the outcomes and events.

Set theory
In a set: $x \in A$
Not in a set: $x \notin A$
Set of tabular method: $S = \{x_1, x_2, x_3, \dots\}$
Set of rule method: $S = \{x | x \text{ satisfies } P\}$
Countable and uncountable are self-explanatory
Empty/null set: $\emptyset = \{\}$
Subsets: $A \subseteq B, A = \{x_1, x_2\}, B = \{x_1, x_2, x_3\}, \emptyset \subseteq A$
Universal set S with N elements, 2^N subsets of S
Union: $A \cup B = \{x | x \in A \text{ or } x \in B\}$
Intersection: $A \cap B = \{x | x \in A \text{ and } x \in B\}$
Complement: $A^c = \{x | x \notin A\}, S^c = \emptyset$
Set difference: $A - B = \{x | x \in A \text{ and } x \notin B\}$
Commutative, distributive, and associative apply to union and intersections.
De Morgan's law: $(A \cup B)^c = A^c \cap B^c, (A \cap B)^c = A^c \cup B^c$
Mutually exclusive (disjoint) sets: $A_i \cap A_j = \emptyset, i \neq j$
Partition: $\cup_{\{i=1\}}^n A_i = A_1 \cup A_2 \cup \dots \cup A_n = S$

Set to probability
Element – Outcome
Set – Event
Universal Set – Sample space

Probability Law: $A \subset S \Longrightarrow^{\{P[\cdot]\}} P[A] \in [0,1] \in \mathbb{R}$
Axioms of Probability:
1. For any event A, $P[A] \geq 0$
2. $P[S] = 1$
3. If $A \cap B = \emptyset$, then $P[A \cup B] = P[A] + P[B]$

Consequences of these axioms:
1. $P[\emptyset] = 0$
2. $P[A^c] = 1 - P[A]$
3. $P[A \cup B] = P[A] + P[B] - P[A \cap B]$
4. If $A \subseteq B$, then $P[A] \leq P[B]$

Union bound: $P[A \cup B] \leq P[A] + P[B]$

More probability
Conditional probability: $P[A|B] \triangleq \frac{P[A \cap B]}{P[B]}$,
$P[A \cap B] = P[A|B]P[B]$
a) $P[A|B] \geq 0$
b) $P[S|B] = 1$
c) If $A = A_1 \cup A_2 \cup \cdots$ with $A_i \cap A_j = \emptyset$ for $i \neq j$, then $P[A|B] = P[A_1|B] + P[A_2|B] + \cdots$

Law of total probability: $P[B] = \sum_{i=1}^n P[B|A_i]P[A_i]$
Bayes' Theorem: $P[A_i|B] = \frac{P[B|A_i]P[A_i]}{P[B]}$
Independence: $P[A \cap B] = P[A]P[B]$
Disjoint: $P[A \cap B] = P[\emptyset] = 0$
INDEPENDENT EVENTS AIN'T DISJOINT AND VICE VERSA
Bernoulli trials: Find probability of an outcome after n trials
Reliability Problem (series): If one fails, whole thing fails
Reliability Problem (parallel): If all fails, whole thing fails

Discrete random variables
Range of an RV: $S_X \subseteq \mathbb{R}$
Discrete RV: $S_X = \{x_z, x_2, \dots\}$
Continuous RV: S_X is an uncountable set
Mixed RV: If it has elements of both

Probability mass function: $p_X(x) = P[X = x]$
1. For any x, $p_X(x) \geq 0$
2. $\sum_{x \in S_X} p_X(x) = 1$
3. For any event $B \subseteq S\_X$, $P[B] = \sum_{x \in B} p_X(x)$

Family of discrete random variables
Bernoulli(p) RV:
- $p_X(x) = \begin{cases} 1-p, & x = 0 \\ p, & x = 1 \\ 0, & \text{otherwise} \end{cases}$
- Used for experiments with 2 outcomes

Binomial(n,p) RV:
- $p_X(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & x = 0,1,2,\dots,n \\ 0, & \text{otherwise} \end{cases}$
- Used in exp. with n trials of 2 outcomes

Poisson(α) RV:
- $p_X(x) = \begin{cases} \frac{\alpha^x e^{-a}}{x!}, & x = 0,1,2,\dots \\ 0, & \text{otherwise} \end{cases}, \alpha = \lambda T$
- Used for number of completely random events

Geometric(p) RV:
- $p_X(x) = \begin{cases} p(1-p)^{x-1}, & x = 1,2,3,\dots \\ 0 & \text{otherwise} \end{cases}$
- Used for no. of tries until desired outcome

Pascal(k,p) RV:
- $p_X(x) = \binom{x-1}{k-1} p^k (1-p)^{x-k}, \ x = k, k+1, \dots$
- Used for probability of k successes

Expected Value
Expected value: $E[X] = \mu_X = \sum x p_X(x)$
Bernoulli: $E[X] = p$
Binomial: $E[X] = np$
Geometric: $E[X] = \frac{1}{p}$
Poisson: $E[X] = \alpha$

Cumulative distribution function
CDF: $F_X(x) = P[X \leq x]$
CDF is cumulative, while PMF is per value
CDF is 1 at the end, while PMF is not
Theorem: $F_X(b) - F_X(a) = P[a < X \leq B]$

Variances for discrete RVs
Bernoulli: $VAR[X] = p(1-p)$
Binomial: $VAR[X] = np(1-p)$
Geometric: $VAR[X] = \frac{1-p}{p^2}$
Poisson: $VAR[X] = \alpha$

Function of RV
Function: $Y = g(X)$
For discrete, PMF is: $p_Y(y) = \sum_{\{x | g(x) = y\}} p_X(x)$
Expected value of Y: $\mu_Y = E[Y] = E[g(x)] = \sum_{x \in S_X} g(x) p_X(x)$
Linear func of RV: $E[aX + b] = aE[X] + b$
Variance: $VAR[X] = \sigma_X^2 \triangleq E[(X - \mu_X)^2] = \sum_{x \in S_X} (x - \mu_X)^2 p_X(x)$
Also variance: $VAR[X] = E[X^2] - (E[X])^2$
Also variance: $VAR[aX + b] = a^2 VAR[X] \in S\_X$
Conditional PMF: $p_{X|B}(x) = P[X = x|B]$

Continuous random variables
For things that are not discrete
CDF of continuous RV: $F_X(x) = P[X \leq x]$
Probability density function (PDF): $f_X(x) = \frac{dF_X(x)}{dx}$
PDF theorem: $P[x_1 < X \leq x_2] = \int_{x_1}^{x_2} f_X(x) dx$
Expected value: $E[X] = \mu_X = \int_{-\infty}^{\infty} x f_X(x) dx, E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$
Variance: $\sigma_X^2 = VAR[X] = E[(X - \mu_X)^2] = \int_{-\infty}^{\infty} (x - \mu_X)^2 f_X(x) dx$

Validity of a PDF: $f_X(x) \geq 0$ and $\int_{-\infty}^{\infty} f_X(x) = 1$

Family of continuous random variables
Uniform(a,b) RV: $f_X(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{otherwise} \end{cases}, F_X(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases}, E[X] = \frac{a+b}{2}, VAR[X] = \frac{(b-a)^2}{12}$

Exponential(λ) RV: $f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}, F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}, E[X] = \frac{1}{\lambda}, \text{VAR}[X] = \frac{1}{\lambda^2}$

Erlang(n,λ): $f_X(x) = \begin{cases} \frac{\lambda e^{-\lambda x} (\lambda x)^{n-1}}{\Gamma(n)} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}, \Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \ (\alpha > 0), E[X] = \frac{n}{\lambda}, VAR[X] = \frac{n}{\lambda^2}$

Gaussian(μ, σ^2) : $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x-\mu)^2/(2\sigma^2)} \ -\infty < x < \infty, E[X] = \mu, VAR[X] = \sigma^2$

CDF of standard normal RV Z: $F_Z(z) = P[Z \leq z] = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du \triangleq \Phi(z), P[X \leq x] = F_X(x) = \Phi(\frac{x-\mu}{\sigma})$

Conditioning continuous RV
Cond. PDF: $f_{X|B}(x) = \begin{cases} \frac{f_X(x)}{P[B]} & x \in B \\ 0 & \text{otherwise} \end{cases}, \text{E}[X|B] = \int_{-\infty}^{\infty} x f_{X|B}(x) dx, VAR[X|B\} = E[X^2|B] - E[X|B]^2$

Memoryless property: Let X be an exponential RV
$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}, F_X(x) = 1 - e^{-\lambda x} \quad x \geq 0$$
Then we have $P[X > x + h | X > x] = P[X > h]$