Probability Spaces

Sample Space Ω => set of **all outcomes** *(mutually exclusive)* of **random experiment**
Event $E \subseteq \Omega$ => any subset of sample space
Extreme Events => **null event** $\emptyset$ & **universal event** S
Elementary Event => *singleton subsets* of S
If $s \in E$ is *experiment outcome*, then E occurred
Null event $\emptyset$ never & *universal event* S always -occurs
For events $E_1, E_2, \dots$:
E_1 *or* E_2 *or* ... => $\bigcup_i E_i$
E_1 *and* E_2 *and* ... => $\bigcap_i E_i$
$E_1, E_2, \dots$ are **Mutually Exclusive** => $\forall i, j.\ E_i \cap E_j = \emptyset$ i.e. they're *pairwise-disjoint*
$E_1, E_2, \dots$ are **Independent Events** =>
$P(\bigcap_{j=1}^n E_{i_j}) = \prod_{j=1}^n P(E_{i_j})$ for any finite subset $\{E_{i_1}, E_{i_2}, \dots, E_{i_n}\}$
If events A, B are *independent*, then $\bar{A}, B$ are *also independent*

σ-algebra $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ => family of subsets of Ω s.t.
nonempty: $\Omega \in \mathcal{F}$
closed under *complements*: $E \in \mathcal{F} \implies \bar{E} \in \mathcal{F}$
closed under *countable union*:
$E_1, E_2, \dots \in \mathcal{F} \implies \bigcup_i E_i \in \mathcal{F}$
Immediate Basic Results:
$\emptyset \in \mathcal{F}$
closed under *countable intersection*:
$E_1, E_2, \dots \in \mathcal{F} \implies \bigcap_i E_i \in \mathcal{F}$
$\{\emptyset, \Omega\}$ is smallest & $\mathcal{P}(\Omega)$ is largest *σ-algebras*
Generated σ-algebras:
$\sigma(\mathcal{G})$ => for family of subsets $\mathcal{G} \subseteq \mathcal{P}(\Omega)$, its *smallest σ-algebra* to contain $\mathcal{G}$ *(exists & unique)*
$\sigma(f)$ => for $f : \Omega \to E$ where $(E, \mathcal{E})$ is *measurable space*,
$\sigma(f) = \{f^{-1}[F] \mid F \in \mathcal{E}\}$ i.e. all pre-images
trace σ-algebra of $B \in \mathcal{F}$ => $\mathcal{F}_B = \{B \cap A \mid A \in \mathcal{F}\}$

Probability Measure $P : \mathcal{F} \to [0, 1]$ on $(\Omega, \mathcal{F})$
$\forall E \in \mathcal{F}, 0 \le P(E) \le 1$ i.e. between 0 and 1
$P(\Omega) = 1$ i.e. *universal event* S always occurs
σ-additive *(countably additive)* => $P(\bigcup_i E_i) = \sum_i P(E_i)$
for *pairwise-disjoint* events $E_1, E_2, \dots \in \mathcal{F}$
Immediate Basic Results:
$P(\bar{E}) = 1 - P(E)$
$P(\emptyset) = 0$
$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Measurable Space $(\Omega, \mathcal{F})$ => *sample space* Ω with *σ-algebra* $\mathcal{F}$ on it
Probability Space $(\Omega, \mathcal{F}, P)$ => *measurable space* $(\Omega, \mathcal{F})$ with *probability measure* P on it

Conditional Probability

Conditional Probability => $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$ where $A, B \subseteq \Omega$ and $P(B) \ne 0$
The **conditional probability space** is $(B, \mathcal{F}_B, P(\cdot \mid B))$
Sample space $B \subseteq \Omega$
Trace σ-algebra $\mathcal{F}_B = \{B \cap A \mid A \in \mathcal{F}\}$
Probability measure $P(\cdot \mid B)$
If A, B are *independent* then $P(A \mid B) = P(A)$
A_1, A_2 are **Conditionally Independent** given B **iff**
$P(A_1 \cap A_2 \mid B) = P(A_1 \mid B) P(A_2 \mid B)$

Law Of Total Probability => for any events $\{B_1, B_2, \dots\}$ which partition Ω, $P(A) = \sum_i P(A \mid B_i) P(B_i) = \sum_i P(A \cap B_i)$
Special Case =>
$P(A) = P(A \cap B) + P(A \cap \bar{B}) = P(A \mid B) P(B) + P(A \mid \bar{B}) P(\bar{B})$

Bayes Theorem => $P(A \mid B) = \frac{P(B \mid A) P(A)}{P(B)}$

General Random Variables

Random Variable => *measurable function* $X : \Omega \to E$
$(\Omega, \mathcal{F}, P)$ is a *probability space*, $(E, \mathcal{E})$ is a *measurable space*
For every $B \in \mathcal{E}$ the pre-image of B under X is in $\mathcal{F}$,
i.e. $X^{-1}[B] = \{s \in \Omega \mid X(s) \in B\} \in \mathcal{F}$
i.e. $\sigma(X) \subseteq \mathcal{F}$ where $\sigma(X)$ is generated by function X
$g(X)(s) = (g \circ X)(s)$ is *also random variable*, for *measurable function* $g : E \to E$
Induced Probability $P_X(X \in B)$ => probability that X takes on value in $B \in \mathcal{E}$
$P_X(X \in B) = P(X^{-1}[B]) = P(\{s \in \Omega \mid X(s) \in B\})$
Also called **Pushforward Measure** of P onto $(E, \mathcal{E})$
induced by X => $(E, \mathcal{E}, P_X)$ is a *probability space*
Also called the **Probability Distribution** of X

Real Random Variables

Real Random Variable is *RV* who's co-domain is $E = \mathbb{R}$
Support $\mathrm{supp}(X)$ => is range of X i.e. $\mathrm{supp}(X) = X[\Omega]$
Simple RRV iff *finite* $\mathrm{supp}(X)$
Discrete RRV iff *countable* $\mathrm{supp}(X)$
Continuous RRV => *uncountable* $\mathrm{supp}(X)$
Induced Probability => $P_X(X \le x) = P(\{s \in \Omega \mid X(s) \le x\})$

Cumulative Distribution Func. (CDF) $F_X(x) = P_X(X \le x)$
$F_X(x)$ is *right-continuous* => for any *decreasing* $\langle x_n \rangle$, $\lim_{n\to\infty} x_n = x_L \implies \lim_{n\to\infty} F_X(x_n) = F_X(x_L)$
To check that function is valid CDF, must obey:
Monotonicity $\forall x_1, x_2 \in \mathbb{R}, x_1 < x_2 \implies F_X(x_1) \le F_X(x_2)$
$F_X(-\infty) = 0, F_X(\infty) = 1$
F_X is right-continuous
Simple Properties
$0 \le F_X(x) \le 1, \forall x \in \mathbb{R}$
$P_X(a < X \le b) = F_X(b) - F_X(a)$ for finite intervals $(a, b] \subseteq \mathbb{R}$

Moments of RRVs

Expectation $E[X]$ => the **mean** μ_X of *distribution* of X
Discrete => $E_X[g(X)] = \sum_x g(x) p_X(x)$
Continuous => $E_X[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$
Linearity => $E[\alpha g(X) + h(X) + \beta] = \alpha E[g(X)] + E[h(X)] + \beta$
Sum => for any $X_1, \dots, X_n$:
$E[\sum_{i=1}^n X_i] = \sum_{i=1}^n E[X_i]$ $E[\bar{X}] = E[\frac{\sum_{i=1}^n X_i}{n}] = \frac{\sum_{i=1}^n E[X_i]}{n}$
Independent Product => $E[\prod_{i=1}^n X_i] = \prod_{i=1}^n E[X_i]$
Independent and Identically Distributed (i.i.d.) =>
$E(\bar{X}) = \mu_X$

n-th Raw Moment $\mu'_n = E[X^n]$ => i.e. about zero
n-th Central Moment $\mu_n = E[(X - E[X])^n]$

Variance $\sigma_X^2 = \mathrm{Var}(X) = E[(X - E(X))^2]$
$\mathrm{Var}(X) = E[X^2] - (E[X])^2$
$\mathrm{Var}(aX + b) = a^2 \mathrm{Var}(X)$
Sum => for *independent* $X_1, \dots, X_n$:
$\mathrm{Var}(\sum_{i=1}^n X_i) = \sum_{i=1}^n \mathrm{Var}(X_i)$
$\mathrm{Var}(\bar{X}) = \mathrm{Var}(\frac{\sum_{i=1}^n X_i}{n^2}) = \frac{\sum_{i=1}^n \mathrm{Var}(X_i)}{n}$
Independent and Identically Distributed (i.i.d.) =>
$\mathrm{Var}(\bar{X}) = \frac{\sigma_X^2}{n}$

Standard Deviation $\sigma_X = \mathrm{sd}(X) = \sqrt{\mathrm{Var}(X)}$

n-th Standardized Moment $\tilde{\mu}_n = \frac{\mu_n}{\sigma^n} = \frac{E[(X - E[X])^n]}{\sqrt{\mathrm{Var}(X)}^n}$

Skewness $\gamma_1 = \tilde{\mu}_3 = \frac{E[(X - \mu)^3]}{\sigma^3}$ *measures asymmetry*
positive skew => distribution *leans left*
negative skew => distribution *leans right*

Moment Generating Function (MGF) $M_X(t) = E[e^{tX}]$
$E[X^n] = \left.\frac{d^n M_X}{dt^n}\right|_{t=0}$ if *open interval* around $t = 0$ exists
because $e^{tX} = 1 + tX + \frac{t^2 X^2}{2!} + \frac{t^3 X^3}{3!} + \dots + \frac{t^n X^n}{n!} + \dots$
so $M_X(t) = E[e^{tX}] = 1 + tE[X] + \frac{t^2 E[X^2]}{2!} + \dots + \frac{t^n E[X^n]}{n!} + \dots$
Sum => for *independent* $X_1, \dots, X_n$
let $X = \sum_{i=1}^n X_i$ => $M_X(t) = \prod_{i=1}^n M_{X_i}(t)$

Discrete Random Variables

Discrete RRV iff *countable* $\mathrm{supp}(X)$
Let $\mathrm{supp}(X) = \{x_1, x_2, \dots\}$ be ordered s.t. $x_1 < x_2 < \dots$
CDF F_X will be *monotonic increasing step function*,
i.e. $F_X(x_i) = F_X(x_{i-1}) + P_X(X = x_i)$
i.e. $P_X(X = x_i) = F_X(x_i) - F_X(x_{i-1})$
Probability Mass Function (PMF) $p(x) = P_X(X = x)$
$0 \le p(x) \le 1, \forall x \in \mathbb{R}$
$\sum_{x \in \mathrm{supp}(X)} p(x) = 1$
$p(x_i) = F(x_i) - F(x_{i-1})$
$F(x_i) = \sum_{j=1}^i p(x_j)$

Bernoulli(P) => TODO: HEREEE!!!!!
Binomial(N,p) => TODO: HEREEE!!!!!
Poisson(P) => TODO: HEREEE!!!!!
Poisson(Λ) => TODO: HEREEE!!!!!
Uniform(U) => TODO: HEREEE!!!!!
Negative Binomial Distribution(U) => TODO: HEREEE!!!!!
Poisson Binomial Distribution => TODO: HEREEE!!!!!

Continuous Random Variables

X is **(Absolutely) Continuous RRV** if $\exists f_X : \mathbb{R} \to \mathbb{R}$ such that $F_X(x) = \int_{u=-\infty}^x f_X(u) du$
f_X called **Probability Density Function (PDF)** of X
$P_X(a < X \le b) = P_X(X \le b) - P_X(X \le a)$
$P_X(a < X \le b) = F_X(b) - F_X(a) = \int_a^b f_X(x) dx$
$P_X(X = x) = P_X(\{x\}) = 0$ and
$P_X(X \in \{x_1, x_2, \dots\}) = P_X(X = x_1) + P_X(X = x_2) + \dots$ for *countable sets*
Properties of PDFs:
$F_X(x) = \int_{-\infty}^x f_X(t) dt$ & $f_X(x) = \frac{d}{dx} F_X(x)$
$f_X(x) \ge 0$ & $\int_{-\infty}^{\infty} f_X(x) dx = 1$
$f_X(x) \cdot h \approx P_X(X \in [x, x + h])$

Quantiles and Percentiles:
The **lower** and **upper quartiles** and **median** of sample of data are the points $(\frac{1}{4}, \frac{3}{4}, \frac{1}{2})$-way through the ordered dataset, respectively
α-quantile $Q_X(\alpha)$ => for continuous X and $0 \le \alpha \le 1$, the least number satisfying $P(X \le Q_X(\alpha)) = \alpha$
i.e. $Q_X(\alpha) = F_X^{-1}(\alpha)$
$Q(1/2)$ is *median* & **k-th percentile** is $Q(k/100)$

Uniform(A,b) => TODO: HEREEE!!!!!
Exp(Λ) => TODO: HEREEE!!!!!
Normal(M, Σ2) => TODO: HEREEE!!!!!
Lognormal => TODO: HEREEE!!!!!

Joint Random Variables

Product Probability Space $(\prod_i \Omega_i, \bigotimes_i \mathcal{F}_i, \bigtimes_i P_i)$
Let $(\Omega_1, \mathcal{F}_1, P_1), \dots, (\Omega_n, \mathcal{F}_n, P_n)$ be *probability spaces*
$\prod_i \Omega_i$ is *n-ary Cartesian product* of $\Omega_1, \dots, \Omega_n$
$\bigotimes_i \mathcal{F}_i$ is *product σ-algebra* of $\mathcal{F}_1, \dots, \mathcal{F}_n$
$\bigotimes_i \mathcal{F}_i \triangleq \sigma(\{\prod_{i=1}^n E_i \mid E_1 \in \mathcal{F}_1, \dots, E_n \in \mathcal{F}_n\})$
where $\sigma(\cdot)$ is the *generated σ-algebra*
$\bigtimes_i P_i$ is *unique measure* such that
$(\bigtimes_i P_i)(E_1, \dots, E_n) = P_1(E_1) \cdots P_n(E_n)$
for every $E_1 \in \mathcal{F}_1, \dots, E_n \in \mathcal{F}_n$

Multivariate Random Variable => $\mathbf{X} : \Omega \to \mathbb{R}^n$
Consider *RRVs* $X_1, \dots, X_n : \Omega \to \mathbb{R}$
NOTE: take *product space if* their *sample spaces differ*
$\mathbf{X} = (X_1, \dots, X_n) : \Omega \to \mathbb{R}^n$ is a **Random Vector**
$[\mathbf{X}(s)]_i = X_i(s)$ => it's a *random variable*
from $(\Omega, \mathcal{F}, P)$ to $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$
$\mathcal{B}(\mathbb{R}^n)$ is **Borel σ-algebra** on $\mathbb{R}^n$
i.e. requirement that each X_i is *also an RV*

Induced Probability =>
$P_{\mathbf{X}}(X_1 \le x_1, \dots, X_n \le x_n)$
$= P(\{s \in \Omega \mid X_1(s) \le x_1 \wedge \dots \wedge X_n(s) \le x_n\})$
RECALL: this *pushforward measure* is *also* called *probability distribution* of X => this *measure* on *Borel σ-algebras* is *also* called the **Joint Probability Distribution**

Joint CDF $F_{\mathbf{X}}(x_1, \dots, x_n) = P_{\mathbf{X}}(X_1 \le x_1, \dots, X_n \le x_n)$
Recover **Marginal CDFs** for each X_i with
$F_{X_i}(x_i) = \lim_{j \ne i, x_j \to \infty} F_{\mathbf{X}}(x_1, \dots, x_n)$
e.g. for *RVs* X, Y => $F_X(x) = F(x, \infty)$ and $F_Y(y) = F(\infty, y)$

To check that function is valid CDF, must obey:
Monotonicity:
For every i, *hold fixed* *all-but-the* i-th component
Then for every x_{i_1}, x_{i_2} we have
$x_{i_1} < x_{i_2} \implies F_{\mathbf{X}}(\dots, x_{i_1}, \dots) \le F_{\mathbf{X}}(\dots, x_{i_2}, \dots)$
$0 \le F_{\mathbf{X}}(x_1, \dots, x_n) \le 1$
$F_{\mathbf{X}}(\dots, -\infty, \dots) = 0, F_{\mathbf{X}}(\infty, \dots, \infty) = 1$
Interval Probabilities for *bi-variate* case $Z = (X, Y)$
$P_Z(x_1 < X \le x_2, Y \le y) = F(x_2, y) - F(x_1, y)$, hence
$P_Z(x_1 < X \le x_2, y_1 < Y \le y_2)$
$= F(x_2, y_2) - F(x_1, y_2) - F(x_2, y_1) + F(x_1, y_1)$

Joint PMF $p(x_1, \dots, x_n) = P_{\mathbf{X}}(X_1 = x_1, \dots, X_n = x_n)$
Recover **Marginal PMFs** for each X_i with
$p_{X_i}(x_i) = \sum_{x_n} \cdots \sum_{x_{i+1}} \sum_{x_{i-1}} \cdots \sum_{x_1} p_{\mathbf{X}}(x_1, \dots, x_n)$
e.g. for *RVs* X, Y => $p_X(x) = \sum_y p(x, y)$ & $p_Y(y) = \sum_x p(x, y)$

To check that function is valid PMF, must obey:
$0 \le p_{\mathbf{X}}(x_1, \dots, x_n) \le 1$ & $\sum_{x_1} \cdots \sum_{x_n} p_{\mathbf{X}}(x_n, \dots, x_1) = 1$

Multinomial Distribution => TODO: HEREEE!!!!!

$X_1, \dots, X_n$ are **Jointly Continuous** if $\exists f_{\mathbf{X}} : \mathbb{R}^n \to \mathbb{R}$ such
$$F_{\mathbf{X}}(x_1, \dots, x_n) = \int_{t_n=-\infty}^{x_n} \cdots \int_{t_1=-\infty}^{x_1} f_{\mathbf{X}}(t_1, \dots, t_n)\ dt_1 \cdots dt_n$$
$f_{\mathbf{X}}$ called **Joint PDF** of $\mathbf{X} = (X_1, \dots, X_n)$
$f_{\mathbf{X}}(x_1, \dots, x_n) = \frac{\partial^n}{\partial x_1 \cdots \partial x_n} F_{\mathbf{X}}(x_1, \dots, x_n)$
To check that function is valid PMF, must obey:
$f_{\mathbf{X}}(x_1, \dots, x_n) \ge 0$ &
$$\int_{t_n=-\infty}^{\infty} \cdots \int_{t_1=-\infty}^{\infty} f_{\mathbf{X}}(t_1, \dots, t_n) dt_1 \cdots dt_n = 1$$
Recover **Marginal PDFs** for each X_i with
$$f_{X_i}(x_i) = \int_{x_n=-\infty}^{\infty} \cdots \int_{x_{i+1}=-\infty}^{\infty} \int_{x_{i-1}=-\infty}^{\infty} \cdots \int_{x_1=-\infty}^{\infty} f_{\mathbf{X}}(x_1, \dots, x_n)\ dx_1 \cdots dx_n$$
e.g. for *RVs* X, Y => $f_X(x) = \int_{y=-\infty}^{\infty} f(x, y) dy$ and
$f_Y(y) = \int_{x=-\infty}^{\infty} f(x, y) dx$

X, Y are **Independent RRVs if-and-only-if**
General => $F(x, y) = F_X(x) F_Y(y)$
Discrete => $p(x, y) = p_X(x) p_Y(y)$
Continuous => $f(x, y) = f_X(x) f_Y(y)$

Joint Expectation $E[\mathbf{X}] = E[X_1, \dots, X_n]$
Discrete =>
$E_{\mathbf{X}}[g(\mathbf{X})] = \sum_{x_n} \cdots \sum_{x_1} g(x_1, \dots, x_n) p_{\mathbf{X}}(x_1, \dots, x_n)$
Continuous =>
$$E_{\mathbf{X}}[g(\mathbf{X})] = \int_{x_n=-\infty}^{\infty} \cdots \int_{x_1=-\infty}^{\infty} g(x_1, \dots, x_n) f_{\mathbf{X}}(x_1, \dots, x_n) dx_1 \cdots dx_n$$
Sum => $E[\sum_i g_i(X_i)] = \sum_i E_{X_i}[g_i(X_i)]$
Independent Product => $E[\prod_i g_i(X_i)] = \prod_i E_{X_i}[g_i(X_i)]$

Covariance =>
$\sigma_{XY} = \mathrm{Cov}(X, Y) = E[(X - \mu_X)(Y - \mu_Y)] = E[XY] - \mu_X \mu_Y$
NOTE: $\mathrm{Var}(X) = \mathrm{Cov}(X, X)$
For *independent RVs* $E[XY] = E_X[X] E_Y[Y]$ so $\sigma_{XY} = 0$

Correlation => $\rho_{XY} = \mathrm{Cor}(X, Y) = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{\mathrm{Cov}(X,Y)}{\mathrm{sd}(X)\,\mathrm{sd}(Y)}$
Correlation is *invariant to the scale* of the X, Y
NOTE: $\mathrm{Var}(X) = \mathrm{Cov}(X, X)$
For *independent RVs* $\sigma_{XY} = \rho_{XY} = 0$

Multivariate Normal Distribution => TODO: HEREEE!!!!!

Conditional Distributions

Conditional Distribution $P_{Y|X}(B_Y \mid B_X) = \frac{P_{XY}(B_X, B_Y)}{P_X(B_X)}$
for any $B_X, B_Y \in \mathbb{R}$
Probability of Y falling inside B_Y given that we know X fell inside B_X

Conditional CDF $F_{Y|X}(y \mid x) = P(Y \le y \mid X = x)$
Discrete => $P(Y \le y \mid X = x) = \sum_{u=-\infty}^{y} p_{Y|X}(u \mid x)$
Continuous => $P(Y \le y \mid X = x) = \int_{u=-\infty}^{y} f_{Y|X}(u \mid x) du$
Interval Probabilities =>
$P(a < X \le b \mid Y = y) = F_{X|Y}(b \mid y) - F_{X|Y}(a \mid y)$
Law of Total Probability CDF =>
$F_X(x) = \int_{y=-\infty}^{\infty} F_{X|Y}(x \mid y) f_Y(y) dy$

Conditional PMF $p_{Y|X}(y \mid x) = \frac{p_{XY}(x,y)}{p_X(x)}$ if $p_X(x) > 0$
Bayes Theorem PMF => $p_{X|Y}(x \mid y) = \frac{p_{Y|X}(y|x) p_X(x)}{p_Y(y)}$
Law of Total Probability PMF =>
$p_X(x) = \sum_y p_{X|Y}(x \mid y) p_Y(y)$

Conditional PDF $f_{Y|X}(y \mid x) = \frac{f_{XY}(x,y)}{f_X(x)}$ if $p_X(x) > 0$
Bayes Theorem PDF => $f_{X|Y}(x \mid y) = \frac{f_{Y|X}(y|x) f_X(x)}{f_Y(y)}$
Law of Total Probability PDF =>
$f_X(x) = \int_{y=-\infty}^{\infty} f_{X|Y}(x \mid y) f_Y(y) dy$

Conditional Expectation $E_{Y|X}[Y \mid X = x]$
Discrete => $E_{Y|X}[Y \mid X = x] = \sum_y y p_{Y|X}(y \mid x)$
Continuous => $E_{Y|X}[Y \mid X = x] = \int_{y=-\infty}^{\infty} y f_{Y|X}(y \mid x) dy$
NOTE: $E_{Y|X}[Y \mid X = \cdot]$ is a function of x

Law of Total Expectation:
Define $E_{Y|X}[Y \mid X](s) = E_{Y|X}[Y \mid X = X(s)]$
Now $E_{Y|X}[Y \mid X] : \Omega \to \mathbb{R}$ is *function of RRV* X
Then we have $E_Y[Y] = E_X[E_{Y|X}[Y \mid X]]$
e.g. in *continuous case* => $\int_x \int_y y f_{Y|X}(y \mid x) f_X(x) dy dx$
$= \int_y \int_x y f(x, y) dx dy$
$= \int_y y f_Y(y) dy$

Markov Chains

Discrete Time Markov Chains (DTMC) => generalization of *sequences of i.i.d. RVs* to support *arbitrary (and possibly dependent) RVs*
State Space J => elements $j \in J$ are *states*
$X_n, n \ge 0$ => takes values in J & models state at time n
Realization of $X_0, X_1, \dots$ is called **Sample Path**
GOAL => calculate $P(X_n = j)$
i.e. probability that at time n system reaches state j
Markov property =>
$P(X_{n+1} = j_{n+1} \mid X_n = j_n, \dots, X_1 = j_1, X_0 = j_0)$
$= P(X_{n+1} = j_{n+1} \mid X_n = j_n)$
i.e. choice of the next state depends on the current state only
How to *specify DTMCs*:
Initial Probability Vector => row-vector $\pi_0^T \in \mathbb{R}^n$
where $(\pi_0)_i = P(X_0 = i)$
Transition Probability Matrix => $R \in \mathbb{R}^{n \times n}$ where
$R_{ij} = P(X_{n+1} = j \mid X_n = i)$
Each transition probability R_{ij} is independent of the time n
Self-loops are allowed => e.g. $R_{ii} = 1$ means DTMC can *never leave* state i *(e.g. a permanent fault)*
R is a *non-nonegative matrix* who's rows sum to 1,
Also called **stochastic matrix**
$P(X_n = j \mid X_0 = i) = (R^n)_{ij}$ => $P(X_n = j) = (\pi_0 R^n)_j$

Classification Of DTMCs:
Irreducible DTMC => *directed graph* associated to R is *strongly connected*:
For any pair (i, j) there *exist some sample path* where, starting in state i, the DTMC eventually reaches state j
Periodic DTMC => *time to return to a state* is an *integer multiple of a fixed period*; otherwise it's **aperiodic DTMC**

Common Long-Term Behaviours Of DTMC:
Steady-State/Stationary Distribution => row-vector π^*_∞ thats invariant under R,
i.e. $P(X_n = j) = (\pi^*_\infty)_j$ for all $n \ge 0, j \in J$
Limiting Distribution => row-vector π_∞ such that
$\pi_\infty = \lim_{n\to+\infty} \pi_0 R^n$
If exits, its *also a steady-state distribution* => converse isn't true
Uniqueness and existence:
The two above are **not unique in general**
If *irreducible and aperiodic DTMC*:
Limiting & steady-state distributions both exist => *unique* and identical to each other
Elements of π_∞ all *strictly positive*
π_∞ is solution of $\pi_\infty R = \pi_\infty$ subject to $\sum_j (\pi_\infty)_j = 1$
i.e. left-eigenvector with left-eigenvalue $\lambda = 1$
If *irreducible but **not** aperiodic DTMC*:
limiting distribution π_∞ no longer exists
Steady-state distribution π^*_∞ still exists =>
its the *unique positive solution* of $\pi^*_\infty R = \pi^*_\infty$ subject to $\sum_j (\pi^*_\infty)_j = 1$

Classification of DTMCs

When can we guarantee existence and uniqueness?

$$R_1 = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{4} & \frac{3}{4} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad R_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \quad R_3 = \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$$

Uniqueness and existence conditions

If a DTMC is irreducible and aperiodic, then:
Without aperiodicity, an irreducible DTMC has .
However, the also in this case and it is the .