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** $\varnothing$ & **universal event** S
- **Elementary Event** => *singleton subsets* of S
- If $s \in E$ is *experiment outcome*, then E **occurred**
  - *Null event* $\varnothing$ never & *universal event* S always -occurs
- For events $E_1, E_2, \ldots$:
  - E_1 *or* E_2 *or …* => $\bigcup_i E_i$
  - E_1 *and* E_2 *and …* => $\bigcap_i E_i$
  - $E_1, E_2, \ldots$ are **Mutually Exclusive** => $\forall i, j.\ E_i \cap E_j = \varnothing$ i.e. they're *pairwise-disjoint*
  - $E_1, E_2, \ldots$ are **Independent** => $\mathrm{P}(\bigcap_{j=1}^{n} E_{i_j}) = \prod_{j=1}^{n} \mathrm{P}(E_{i_j})$ for any finite subset $\{E_{i_1}, E_{i_2}, \ldots, E_{i_n}\}$
    - 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*: $E_1, E_2, \cdots \in \mathcal{F} \implies \bigcup_i E_i \in \mathcal{F}$
- **Immediate Basic Results**:
  - $\varnothing \in \mathcal{F}$
  - closed under *countable intersection*: $E_1, E_2, \cdots \in \mathcal{F} \implies \bigcap_i E_i \in \mathcal{F}$
  - $\{\varnothing, \Omega\}$ 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 $\mathrm{P} : \mathcal{F} \to [0, 1]$ on $(\Omega, \mathcal{F})$

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

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

Probability Space $(\Omega, \mathcal{F}, \mathrm{P})$ => *measurable space* $(\Omega, \mathcal{F})$ with *probability measure* P on it

Conditional Probability

Conditional Probability => $\mathrm{P}(A \mid B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$ where $A, B \subseteq \Omega$ and $\mathrm{P}(B) \neq 0$

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

Law Of Total Probability => for any events $\{B_1, B_2, \ldots\}$ which partition Ω, $\mathrm{P}(A) = \sum_i \mathrm{P}(A \mid B_i)\mathrm{P}(B_i) = \sum_i \mathrm{P}(A \cap B_i)$

- *Special Case* => $\mathrm{P}(A) = \mathrm{P}(A \cap B) + \mathrm{P}(A \cap \bar{B}) = \mathrm{P}(A \mid B)\mathrm{P}(B) + \mathrm{P}(A \mid \bar{B})\mathrm{P}(\bar{B})$

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

General Random Variables

Random Variable => *measurable function* $X : \Omega \to E$

- $(\Omega, \mathcal{F}, \mathrm{P})$ 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 $\mathrm{P}_X(X \in B)$ => probability that X takes on value in $B \in \mathcal{E}$

- $\mathrm{P}_X(X \in B) = \mathrm{P}(X^{-1}[B]) = \mathrm{P}(\{s \in \Omega \mid X(s) \in B\})$
- Also called **Pushforward Measure** of P onto $(E, \mathcal{E})$ induced by X => $(E, \mathcal{E}, \mathrm{P}_X)$ 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** => $\mathrm{P}_X(X \le x) = \mathrm{P}(\{s \in \Omega \mid X(s) \le x\})$

Cumulative Distribution Func. (CDF) $F_X(x) = \mathrm{P}_X(X \le x)$

- $F_X(x)$ 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**
  - $0 \le F_X(x) \le 1, \forall x \in \mathbb{R}$
  - $\mathrm{P}_X(a < X \le b) = F_X(b) - F_X(a)$ for finite intervals $(a, b] \subseteq \mathbb{R}$

Moments of RRVs

Expectation $\mathrm{E}[X]$ => the **mean** μ_X of *distribution* of X

- *Discrete* => $\mathrm{E}[g(X)] = \sum_x g(x)p(x)$
- *Continuous* => TODO: HEREEE!!!!!
- **Linearity** => $\mathrm{E}[\alpha g(X) + h(X) + \beta] = \alpha \mathrm{E}[g(X)] + \mathrm{E}[h(X)] + \beta$
- **Sum** => for any $X_1, \ldots, X_n$:
  - $\mathrm{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathrm{E}[X_i]$ $\mathrm{E}[\bar{X}] = \mathrm{E}[\frac{\sum_{i=1}^{n} X_i}{n}] = \frac{\sum_{i=1}^{n} \mathrm{E}[X_i]}{n}$
  - **Independent and Identically Distributed (i.i.d.)** => $E(\bar{X}) = \mu_X$

n-th Raw Moment $\mu'_n\ \mathrm{E}[X^n]$ => i.e. about zero

n-th Central Moment $\mu_n = \mathrm{E}[(X - E[X])^n]$

Variance $\sigma_X^2 = \mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}(X))^2]$

- $\mathrm{Var}(X) = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$
- $\mathrm{Var}(aX + b) = a^2\ \mathrm{Var}(X)$
- **Sum** => for *independent* $X_1, \ldots, X_n$:
  - $\mathrm{Var}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$
  - $\mathrm{Var}(\bar{X}) = \mathrm{Var}(\frac{\sum_{i=1}^{n} X_i}{n^2}) = \frac{\sum_{i=1}^{n} \mathrm{Var}(X_i)}{n}$
  - **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 $\tilde{\mu}_n = \frac{\mu_n}{\sigma^n} = \frac{\mathrm{E}[(X - E[X])^n]}{\sqrt{\mathrm{Var}(X)}^n}$

Skewness $\gamma_1 = \tilde{\mu}_3 = \frac{\mathrm{E}[(X-\mu)^3]}{\sigma^3}$ *measures asymmetry*

- positive skew => distribution *leans left*
- negative skew => distribution *leans right*

Discrete Random Variables

Discrete RRV **iff** *countable* $\mathrm{supp}(X)$

- Let $\mathrm{supp}(X) = \{x_1, x_2, \ldots\}$ be ordered s.t. $x_1 < x_2 < \ldots$
- *CDF* F_X will be *monotonic increasing step function*, i.e. $F_X(x_i) = F_X(x_{i-1}) + \mathrm{P}_X(X = x_i)$ i.e. $\mathrm{P}_X(X = x_i) = F_X(x_i) - F_X(x_{i-1})$

Probability Mass Function (PMF) $p(x) = \mathrm{P}_X(X = x)$

- $0 \le p(x) \le 1, \forall x \in \mathbb{R}$
- $\sum_{x \in \mathrm{supp}(X)} p(x) = 1$
- $p(x_i) = F(x_i) - F(x_{i-1})$
- $F(x_i) = \sum_{j=1}^{i} p(x_j)$

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!!!!!