1 Outcomes, events, and Probability

Definition 1.1. (Random) Experiment is a mechanism that results in random outcomes

Definition 1.2. Sample Space (Ω) is the set of all possible outcomes from and experiment

Definition 1.3. Event is a subset of Ω

Definition 1.4. Mutually Exclusive (Disjoint): $A \cup B = \{\} = \varnothing$

Definition 1.5. Commutative: $A \cup B = B \cup A$ | **Associative**: $(A \cup B) \cup C = A \cup (B \cup C)$ |
Distributive $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$, $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Definition 1.6. A implies B $A \cap B = A$, $A \subset B$

Definition 1.7. Probability Function P on finite Ω assigns each event A a $P(A)$ s.t. (Axioms)
i) $P(A) \geq 0$ ii)$P(\Omega) = 1$ iii)$P(A \cup B) = P(A) + P(B)$ if disjoint.

Definition 1.8. Useful Probabilities: *Union*: $P(A \cup B) = P(A \cap B^c) + P(A \cap B) + P(A^c \cap B)$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
Complement $P(\Omega) = P(A) + P(A^c)$, $P(A^c) = 1 - P(A)$

Definition 1.9. Calculating by counting applies only when i) all outcomes are equally likely ii) Ω is finite then
$P(A) = \frac{\text{number of outcomes belonging to } A}{\text{total number of outcomes in } \Omega}$

Definition 1.10. Product of Sample Space in general is $\Omega = \Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$

Definition 1.11. Permutation of a set $\boldsymbol{N}$ of size $\boldsymbol{n}$ is ${}_N P_n = \frac{N!}{(N-n)!}$.

Definition 1.12. Combination of a set $\boldsymbol{N}$ of size $\boldsymbol{n}$ is $\binom{N}{n} = \frac{N!}{(N-n)! \cdot N!}$

Definition 1.13. De Morgan's Law for A, B, we have $(A \cup B)^c = A^c \cap B^c$ and $(A \cap B)^c = A^c \cup B^c$

2 Conditional Probability

Definition 2.1. Conditional Probability of A given C is, in general, $P(A|C) = \frac{P(A \cap C)}{C}$ for any $P(C) > 0$

Definition 2.2. Law of Total Probability states, for disjoint $C_1, \ldots, C_m$, $C_1 \cup \cdots \cup C_m = \Omega$, $P(A) = \sum_{i=1}^{m} [P(A|C_i) P(C_i)]$

Definition 2.3. Bayes' Rule states, for disjoint $C_1, \ldots, C_m$, $C_1 \cup \cdots \cup C_m = \Omega$, $P_(C_i|A) = \frac{P(A|C_i) \cdot P(C_i)}{\sum_{i=1}^{m} [P(A|C_i) P(C_i)}$

Definition 2.4. Independence between A and B implies $P(A|B) = P(A)$

Definition 2.5. Relation between Independent Probabilities: i) A *independent of* $B \iff A^c$ *independent of* B and
ii) A *independent of* $B \iff B$ *independent of* A

Definition 2.6. Independence of Two or More Events: $P(A_1 \cap A_2 \cap \cdots \cap A_n) = \prod_{i=1}^{n} A_i$

3 Random Variables

Definition 3.1. The **probability mass function** of drv $\mathbf{X}$ is $p : \mathbb{R} \to [0, 1]$ defined by $p(k) = P(x = k)$ for $-\infty < k < \infty$.

Definition 3.2. Discrete Random Variable X takes a countable number of values.

Definition 3.3. Continuous Random Variable X has $P(A \leq X \leq B) = \int_a^b f(x)dx$ with i) $\forall x, f(x) \geq 0$ ii) $\int_{-\infty}^{\infty} f(x)dx = 1$.
f is the **probability density function** of X and $f(x)$ is the **probability density** of X at x. Note: a random variable is *continous* if its cdf is continuous everywhere

Definition 3.4. The **cumulative distribution function** F of a drv or crv $\mathbf{X}$ is the function $F : \mathbb{R} \to [0, 1]$ defined by $F(a) = P(X \leq a)$ for $-\infty < a < \infty$

4 Common Distributions

4.1 Discrete

Definition 4.1. Bernoulli $X \sim$ **Ber**(θ): Parameter θ, $0 \leq \theta \leq 1$ and **pmf** given by $p_X(X) = \begin{cases} \theta & x = 1 \\ 1 - \theta & x = 0 \end{cases}$

Useful for modelling experiments with exactly two possible outcomes

Definition 4.2. Binomial $\text{Bin}(n, \theta)$: Parameters n and θ with $n \in \mathbb{N}$ and $0 \leq \theta \leq 1$ and **pmf** given by $p_X(x) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}$
This distribution describes a sum of n ***independent*** and ***identical*** Bernoulli trials

Definition 4.3. Geometric $\text{Geo}(\theta)$: Parameters θ with $0 \leq \theta \leq 1$ and **pmf** given by $p_X(x) = (1 - \theta)^{x-1} \theta$ for $x \in \mathbb{N}$.
The number of identical Bernoulli trials until the first success

Definition 4.4. Poisson Pois(λ): Parameters λ with $\lambda > 0$ and **pmf** given by $p_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}$ for $x \in \mathbb{Z}^+$
Captures the count of events in a fixed interval of Poisson Process. Assumptions: i) the expected rate λ is constant ii) all events are independent of each other iii) events cannot occur simultaneously

4.2 Continuous

Definition 4.5. Uniform $U(\alpha, \beta)$: on interval $[\alpha, \beta]$ and **pdf** given by $f(x) = \begin{cases} \frac{1}{\beta-\alpha} & \alpha \leq x \leq \beta \\ 0 & \textbf{otherwise} \end{cases}$

Uniform distribution assigns equal probabilities across a fixed interval, useful for modelling *completely arbitrary* experiments/*complete ignorance* about probabilities

Definition 4.6. Exponential Exp(λ): Parameter λ with $\lambda > 0$ and **pdf** $f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & otherise \end{cases}$

Useful for modelling time until next event in a Poisson process. λ is the expected rate of events. Note: F_Y is continuous everywhere, f_Y is discontinuous at 0.
CDF given by $1 - e^{-\lambda x}$

Definition 4.7. Gamma $G \sim$ **Gamma**(α, β): Parameters α, β with $\beta, \alpha > 0$ with **pdf** given by $f(x) = \frac{1}{\Gamma(a)}\beta^\alpha x^{\alpha-1}e^{-\beta x}$ for $x > 0$

Definition 4.8. Normal $N(\mu, \sigma^2)$: Parameters μ, σ^2 with $\sigma^0 > 0$ with **pdf** $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\{-\frac{1}{2}(\frac{x-\mu}{\sigma}^2)\}$
Standard Normal Distribution is when $\mu = 0$ and $\sigma^2 = 1$ with **pdf** $\phi(z) = \frac{1}{\sqrt{2pi}}e^{-\frac{1}{2}z^2}$ and **cdf** $\Phi(a) = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}z^2}dz$
Often used to model observational errors. Note: Normal distributions have symmetry along its centre, μ and σ controls the spread of distribution. Note: $\Phi(-a) = 1 - \Phi(a)$
Note: transform $X \sim N(\mu, \sigma^2)$ to $Z = \frac{X-\mu}{\sigma}$ to use with LUT. $1 - \Phi(a) = P(Z \geq a)$ for $N(0, 1)$

5 Quantile, Percent, and Median

Definition 5.1. Quantile Function of random variable X with **cdf** F is $F^{-1} = \min\{x : F(x) \geq T\}$ for $0 \leq T \leq 1$

6 Useful Stuff

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{1}$$

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{2}$$