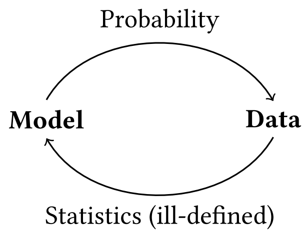
Probability and Statistics

Matthew Scicluna
matthew.scicluna@umontreal.ca

Probability Vs Statistics

Probability: Given Model, how likely is Data? $\rightarrow$ Well-formed since these are Mathematical questions.
Statistics: Given Data, how likely is Model? $\rightarrow$ Ill-formed since many Models can create the same data!



Probability Space

Probability Space: a triple (Ω, F, P) consisting of:

1. Ω the **Sample Space**
2. $F \subseteq 2^{\Omega}$ a σ**-algebra** on Ω i.e.
   (a) $\Omega \in F$
   (b) $E \in F \Rightarrow E^{\complement} \in F$
   (c) $E_1, E_2, \dots \in F \Rightarrow \bigcup_{i=1}^{\infty} E_i \in F$
3. P: $F \mapsto [0, 1]$ a **Probability Measure** i.e.
   (a) $P(E) \geq 0$ for $E \in F$
   (b) $P(\Omega) = 1$
   (c) $P\left(\bigsqcup_{i=1}^{\infty} E_i\right) \Rightarrow \sum_{i=1}^{\infty} P(E_i)$ for $E_i \in F$

Given **Events** $E_i, E \in F$, P also satisfies:

1. **Upward and Downward continuity** of P:
   (a) $E_i \uparrow E \Rightarrow \lim_{n \to \infty} P(E_n) = P(E)$
   (b) $E_i \downarrow E \Rightarrow \lim_{n \to \infty} P(E_n) = P(E)$
2. **Monotonicity** of P:
   (a) $E_i \subseteq E_j \Rightarrow P(E_i) \leq P(E_j)$

Conditional Probability

We can compute Probabilities of Events Conditioned on other Events.
Conditional Probability of event A on event B with $P(B) > 0$ is:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B \mid A)P(A)}{P(B)}$$

A set of events $\{A_i\}$ are **Mutually Independent** if, for any n element subset of $\{A_i\}$:

$$P\left(\bigcap_{i=1}^{n} A_i\right) = \prod_{i=1}^{n} P(A_i)$$

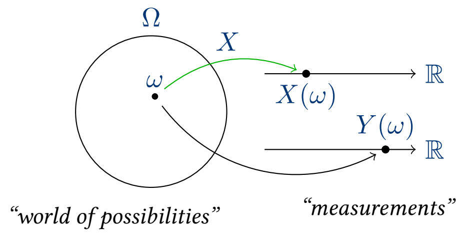
Law of Total Probability: Given Events A and **Partition** $\{B_i\}$ (i.e. where $\bigsqcup_{i=1}^{\infty} B_i = \Omega$)

$$P(A) = \sum_{i=1}^{\infty} P(A \mid B_i)P(B_i)$$

Random Variables

A **Random Variable** is a $\mathbb{B}$-Measurable function
X: $(\Omega, F) \mapsto (\mathbb{R}, \mathbb{B})$

1. For $A \in \mathbb{B}$ we can compute $P(X \in A)$
2. $P(X \in A) := P(X^{-1}(A)) = P(\{\omega \in \Omega : X(\omega) \in A\})$
3. $P(X^{-1}(\cdot)) := P_X(\cdot)$ which is called the **Push-Forward Measure** of P by X on $\mathbb{R}$
4. Hence X induces a new Probability Space $(\mathbb{R}, \mathbb{B}, P_X)$ from the original (Ω, F, P)



Random Variables can be uniquely determined by their **CDF**:
$F_X(t) := P_X((-\infty, t]) = P(X \leq t)$

1. Right Continuous
2. Non-Negative
3. $\lim_{t \to \infty} F_X(t) = 1, \lim_{t \to -\infty} F_X(t) = 0$
4. Any function satisfying above properties is the CDF for some random variable.

Random Variables studied are usually either **Continuous** or **Discrete** (although they can be neither.)

1. If F_X **Absolutely Continuous** then X is Continuous.
   (a) **Absolutely Continuous**: F differentiable a.e. and $\exists f(x)$ s.t. $F_X(x) = \int_{-\infty}^{x} f(u)du$
   (b) f is called the PDF
   (c) Where f is continuous, $\frac{d}{dx} F_X(x) = f(x)$
   (d) hence, f is unique a.e. (may not be everywhere!)
2. If $Range(X)$ is countable then X is Discrete.
3. If X Continuous and Non-negative: **Hazard** of X is $\lambda(t) = \frac{f(t)}{1-F(t)}$
   (a) $1 - F(t) = \exp\left(-\int_0^t \lambda(x)\, dx\right)$
   (b) $\lambda(t)$ interpreted as instantaneous survival rate at time t.
   (c) $\lambda(t) = c \; \forall t \iff X \sim Exp(c)$

Random Vectors

Moments of a Random Variable

Convergence Theorems