Sample Space and Events

The **sample point** of a random experiement is a outcomes that cannot be decomposed into other results. The **sample space** is defiend as all the possible outcomes. IF a sample space is countable, ie the outcomescan be put into a one-to-one correspondence with the positive integers, then it is a **discrete sample space** if it cannot, then it is a **continuous sample space**. An **event** is a subset of the sample space. If the event is the entire sample space then it is a **certain event** if it contains no sample points than it is **impossible** or **null event**. If it contains only one outcome then it is an **elementary event**.

Set Operations

We have that the **union** of two events A and B, $A \cup B$ is the event that either A or B or both occur. The **intersection** of two events A and B, $A \cap B$ is the event that both A and B occur. The **complement** of an event A, A^c is the event that A does not occur. If A is a subset of B then we can say that A **implies** B. And we can say that two events are **equal** if they contain the same outcomes. Some following set properties are usefull

$\boxed{A \cup B = B \cup A}$ $\boxed{A \cap B = B \cap A}$ $\boxed{A \cup (B \cup C) = (A \cup B) \cup C}$ $\boxed{A \cap (B \cap C) = (A \cap B) \cap C}$ $\boxed{(A \cap B)^c = A^c \cup B^c}$

$\boxed{(A \cup B)^c = A^c \cap B^c}$ $\boxed{A \cup (B \cap C) = (A \cup B) \cap (A \cup C)}$ $\boxed{A \cap (B \cup C) = (A \cap B) \cup (A \cap C)}$

Axioms of Probability

Let a random experiment have a sample space S, A probability law for the experiment is a rule that assigns for each event A a number P[A] called the **probability** of A. **Axiom 1:** $P[A] \geq 0$ for all events A. **Axiom 2:** $P[S] = 1$. **Axiom 3:** If $A_1, A_2, \dots$ are disjoint events, ie $A_1 \cap A_2 = \emptyset$ then $P[A_1 \cup A_2] = P[A_1] + P[A_2]$. A generalization of this is that given a family of disjoint events $A_1, A_2, \dots$ then $P[\cup_{i=1}^n A_i] = \sum_{i=1}^n P[A_i]$.

Combinatorics

The **binomial coeeficent** is the number of ways to choose an unordered set of k elements from a set of n elements without replacement. We have that $\binom{n}{k} = \frac{n!}{k!(n-k)!}$. If we do have **replacement** the number of ways for choosing an unordered set of k elements from a set of n elements is $\binom{n-k+1}{k}$. The binomial coefficent is a special case of the **multinomial coefficent** Let us suppose we are partitioning a set of n distinct objects into $\mathcal{F}$ subsets, $B_1, B_2, \dots, B_{\mathcal{F}}$, where B_i has k_i elements and $\sum_{i=1}^{\mathcal{F}} k_i = n$. Then the number of ways to do this is this is $\binom{n}{k_1, k_2, \dots, k_{\mathcal{F}}} = \frac{n!}{k_1! k_2! \cdots k_{\mathcal{F}}!}$.

Conditional Probability and its applications

The **conditional probability** of an event A given B is defined as $P[A|B] = \frac{P[A \cap B]}{P[B]}$. From this we can get the **chain rule** $P[A_1 \cap A_2 \cap \dots \cap A_n] = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1})$. The **total probability rule** is $P[A] = \sum_{i=1}^n P[A|B_i]P[B_i]$ for some $B_1, ...B_n$ that form a partition of the sample space (ie $B_1 \cap B_2 = \emptyset$ and $\cup_{i=1}^n B_i = S$). Therefore from the definition of conditional probability we can derive the bayes rule: $P[B_i|A] = \frac{P[A|B_i]P[B_i]}{P(A)}$

Independence

We say that two events A and B are **independent** if $P[A|B] = P[A]$. From the definition of conditional probability we can derive that independence can be also expressed as $P[A \cap B] = P[A]P[B]$. Furthermore let two events A and B be **conditionally independent** given C if $P[A \cap B|C] = P[A|C]P[B|C]$.

Random Variables

A random variable is a mapping from $(\Omega, \mathcal{F}, P)$ to a measurable space $(\Omega', \mathcal{F}')$, where Σ is the sample space, $\mathcal{F} = 2^\Omega$ is the set of events and P is the probability measure ie $P : \mathcal{F} \to [0, 1]$. More formally a random variable $X : \Omega \to \Omega'$ is a mapping that satisfies: $X^{-1}(B) = \{\omega' \in \Omega : X(\omega') = B\} \in \mathcal{F}$ for all $B \in \mathcal{F}'$.

Discrete Random Variable

In particular, a discrete random variable is a real-valued function of the outcome of the experiment that can take a finite or countably infinite number of values. It is described by its **PMF** which gives the probability of each numerical value that the random variable can take. Furthermore the function of a discrete random variable defines another discrete random variable.

Expectation and Variance

The **expectated value** of a random variable X is defined as $E[X] = \sum_{x \in \Omega'} xP[X = x]$. More generally for a function of a random variable $g(X)$ we have that $E[g(X)] = \sum_{x \in \Omega'} g(x)P[X = x]$. The **variance** of a random variable X is defined as $Var(X) = E[(X - E[X])^2] = E[X^2] - E[X]^2$. If $Y = aX + b$ we have that $E[Y] = aE[X] + b$ and $Var(Y) = a^2 Var(X)$. For two random variables A and B we have $E[A + B] = E[A] + E[B]$ and $Var(A + B) = Var(A) + Var(B) + 2Cov(A, B)$, therefore if A and B are independent $Var(A + B) = Var(A) + Var(B)$ and $E[AB] = E[A]E[B]$.

Common Discrete Random Variables

Bernoulli	Binomial	Geometric	Negative Binomial
$X \sim Ber(p)$	$X \sim Bin(n, p)$	$X \sim Geo(p)$	$X \sim NegBin(r, p)$
$X \in \{0, 1\}$	$X \in \{0, 1, \dots, n\}$	$X \in \{0, 1, \dots\}$	$X \in \{0, 1, \dots\}$
$P[X = 1] = p$	$P[X = k] = \binom{n}{k} p^k (1 - p)^{n-k}$	$P[X = k] = p(1 - p)^k$	$P[X = k] = \binom{k+r-1}{k} p^k (1 - p)^r$
$E[X] = p$	$E[X] = np$	$E[X] = \frac{1}{p}$	$E[X] = \frac{rp}{1-p}$
$Var[X] = p(1 - p)$	$Var[X] = np(1 - p)$	$Var[X] = \frac{1-p}{p^2}$	$Var[X] = \frac{rp}{(1-p)^2}$

Poisson	Hypergeometric	Discrete Uniform	zipf
$X \sim Pois(\lambda)$	$X \sim Hype(n, N, K)$	$X \sim DisUnif(a, b)$	$X \sim Zipf(s, N)$
$X \in \{0, 1, \dots\}$	$X \in \{0, 1, \dots, K\}$	$X \in \{a, a + 1, \dots, b\}$	$X \in \{1, 2, \dots N\}$
$P[X = k] = \frac{\lambda^k e^{-\lambda}}{k!}$	$P[X = k] = \frac{\binom{K}{k}\binom{n-K}{n-k}}{\binom{n}{k}}$	$P[X = a] = \frac{1}{b-a+1}$	$P[X = k] = \frac{1}{H_{N,s}} \frac{1}{k^s}$
$E[X] = \lambda$	$E[X] = n\frac{K}{N}$	$E[X] = \frac{a+b}{2}$	$E[X] = \frac{H_{N,s-1}}{H_{N,s}}$
$Var[X] = \lambda$	$Var[X] = \frac{nK(N-K)(N-n)}{N^2(N-1)}$	$Var[X] = \frac{(b-a+1)^2-1}{12}$	$Var[X] = \frac{H_{N,s-2}}{H_{N,s}} - \frac{H^2_{N,s-1}}{H^2_{N,s}}$

Where $H_{N,s} = \sum_{k=1}^N \frac{1}{k^s}$