Probability

Donald Kong

dkkl1137@gmail.com

:D

Contents

1 Combinatorial Analysis

Theorem (Generalised Principle of Counting). *If r experiments are such that each i-th experiment has n_i possible outcomes, there are a total of $n_1 \cdot n_2 \cdots n_r$ possible outcomes for the r experiments.*

Proof. Enumerating all the possible outcomes of two experiments, we have

$$\begin{array}{cccc} (1,1), & (1,2) & \cdots & (1,n) \\ (2,1), & (2,2) & \cdots & (2,n) \\ \vdots & \vdots & \ddots & \vdots \\ (m,1), & (m,2) & \cdots & (m,n) \end{array}$$

where the outcome is (i,j) if experiment 1 results in the i-th possible outcome and experiment 2 then results in its j-th possible outcome. The set of possible outcomes consists m rows of n elements. Applying this result repeatedly, we extend the principle to r experiments. ∎

Permutations

Proposition. *Suppose we have n objects. There are $n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1 = n!$ distinct permutations of the objects.*

Proposition (Permutations of alike objects). *For n objects, of which n_1 are alike, n_2 are alike, $\ldots$,n_r are alike, we have*

$$\frac{n!}{n_1!\, n_2! \cdots n_r!}$$

distinct permutations.

Combinations

Proposition. *Suppose we have n objects. There are $\frac{n!}{(n-r)!r!}$ different subsets of r objects from the set of n objects.*

Proof. From a set of n objects, we have n possible ways to choose the first object, $(n-1)$ possible ways to choose the second object, $\ldots$, $(n-r+1)$ possible ways to choose the r-th object. By the generalised principle of counting, we have

$$n(n-1)\cdots(n-r+1)$$

ways to select a group of r objects from n objects, when order is relevant. Each group of r objects will be counted $r!$ times in this count, since there are $r!$ permutations of r objects. It follows that there are

$$\frac{n!}{(n-r)!\, r!}$$

distinct groups where order is irrelevant. ∎

Notation ($\binom{n}{r}$). We define $\binom{n}{r}$ for $r \leq A$ as the number of subsets of size r that can be chosen from a set of size n

$$\binom{n}{r} = \frac{n!}{(n-r)!\, r!}.$$

Theorem (Pascal's Identity)**.**

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}, 1 \leq r \leq n$$

Proof. Suppose a group of n objects. Consider a particular object, Object 1. There are $\binom{n-1}{r-1}$ r-sized groups that contain Object 1, and $\binom{n-1}{r}$ r-sized groups that do not contain Object 1. Since there are a total of $\binom{n}{r}$ r-sized groups, Pascal's Identity follows. ■

Theorem (Binomial Theorem)**.**

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}.$$

Proof. Consider the product

$$(x_1 + y_1)(x_2 + y_2) \cdots (x_n + y_n)$$

The expansion consists a sum of 2^n terms, each term a product of n factors, where for each $i = 1, 2, \ldots, n$, either x_i or y_i is contained as a factor. There are $\binom{n}{k}$ terms that have k of the x's and $(n-k)$ of the y's as a factor. Letting $x_i = x, y_i = y$, we have the result.

Notation (Multinomial Coefficients). For $n_1 + n_2 + \cdots + n_r = n$,

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1!\, n_2! \cdots n_r!}.$$

$\binom{n}{n_1, n_2, \ldots, n_r}$ represents the number of possible divisions of n distinct objects into r distinct groups sized $n_1, n_2, \ldots, n_r$.

Theorem (Multinomial Theorem)**.** *Generalisation of the Binomial Theorem.*

$$(x_1 + x_2 + \cdots + x_r)^n = \sum_{\{(n_1, \ldots, n_r) \in \mathbb{N}^r : n_1 + \cdots + n_r = n\}} \binom{n}{n_1, n_2, \ldots, n_r} x_1^{n_1} x_2^{n_2} \ldots x_r^{n_r}.$$

Proposition (Number of Integer Solutions)**.** *For the equation:*

$$x_1 + x_2 + \cdots + x_r = n,$$

there are $\binom{n-1}{r-1}$ distinct positive vectors $(x_1, x_2, \ldots, x_r) \in \mathbb{N}_*^r$ and $\binom{n+r-1}{r-1}$ distinct non-negative vectors $(x_1, x_2, \ldots, x_r) \in \mathbb{N}^r$.

The number of nonnegative solutions of $x_1 + x_2 + \cdots + x_r = n$ equals the number of positive solutions of $y_1 + y_2 + \cdots + y_r = n + r$, by letting $y_i = x_i + 1$.

2 Axioms of Probability

2.1 Sample Space, Events, Experiments

Definition (Experiment)**.** An experiment is any procedure that can be infinitely repeated and has a well-defined set of possible outcomes. It is modeled by a probability space $(S, \mathcal{F}, P)$, consisting a sample space S, an event space $\mathcal{F}$, a probability function P.

Definition (Sample Space)**.** The set of all possible outcomes of an experiment, denoted by S.

Definition (Event Space)**.** A set consisting subsets of S, $\mathcal{F} \subseteq \mathcal{P}(S)$, such that it contains the sample space and is closed under complements and countable unions. Also known as a σ-algebra.

Union, intersection, complement operations are defined on the event space.

Definition (Event)**.** Any $E \in \mathcal{F}$ is known as an event, a subset of S that consists possible outcomes of the experiment. If the outcome of an experiment is in E, we say that E has occurred.

If events $E_1 \cap E_2 \cap \ldots \cap E_n = \varnothing$, we say that $E_1, E_2, \ldots, E_n$ are mutually exclusive events.

Theorem (De Morgan's Laws)**.**

$$\left(\bigcup_{i=1}^{n} E_i\right)^c = \bigcap_{i=1}^{n} E_i{}^c \qquad \left(\bigcap_{i=1}^{n} E_i\right)^c = \bigcup_{i=1}^{n} E_i{}^c.$$

2.2 Kolmogorov's Axioms

$\forall E \in \mathcal{F}, P(E) \in \mathbb{R}$ and satisfies the following:

1. $0 \leq P(E) \leq 1$
2. $P(S) = 1$
3. For any sequence of mutually exclusive events $E_1, E_2, \ldots,$

$$P(\bigcup_{i=1}^{\infty} E_i) = \sum_{i=1}^{\infty} P(E_i)$$

$P(\cdot)$ satisfying the three axioms is known as a probability function.

Remark. For sample spaces that are uncountably infinite sets, $P(E)$ is defined only for *measurable* events. Events of practical interest are measurable.

2.3 Probabilities

Proposition (Results of Kolmogorov's Axioms)**.** *For any events E, F,*

$$P(E^c) = 1 - P(E).$$

$$E \subseteq F \implies P(E) \leq P(F).$$

Theorem (Principle of Inclusion-Exclusion (PIE))**.** *For any events $E_1, E_2, \ldots, E_n$,*

$$\begin{aligned} P(E_1 \cup E_2 \cup \cdots \cup E_n) = &\sum_{i=1}^{n} P(E_i) - \sum_{i_1 < i_2} P(E_{i_1} E_{i_2}) + \cdots \\ &+ (-1)^{r+1} \sum_{i_1 < i_2 < \cdots < i_r} P(E_{i_1} E_{i_2} \cdots E_{i_r}) \\ &+ \cdots + (-1)^{n+1} P(E_1 E_2 \cdots E_n) \end{aligned}$$

Corollary. *Given any two events E, F,*

$$P(E \cup F) = P(E) + P(F) - P(EF).$$