1. Algebra of Events

Q.1.5.1.1 What is the sample space for measuring the life-time of a bulb?
Answer: The life of a bulb can continue indefinitely. Thus, the sample space, $S = [0, \infty)$, where a semi-closed set $[a, b) = \{x : a \leq x < b\}$

Q.1.5.1.2 What do you mean by an event? What is an elementary event?
Answer: An event is a subset of the sample space. An event E is said to occur on a particular trial of the experiment if the outcome observed is an element of the set E. Consider the experiment of flipping two coins. Let X be the event of getting atleast one head. Then $X = \{(H, H), (H, T), (T, H)\}$. Every sample point in the sample space denotes an elementary event. Thus, $Y = \{(H, T)\}$ is an elementary event.

Q.1.5.1.3 Construct the sample space of three-child family that describes the genders of the children with respect to birth order.
Answer: We assume that there are two outcomes at every birth: boy (b) and girl (g). Fig. 1.1 shows a tree structure for a three-child family. Sample space, $\Omega = \{bbb, bbg, bgb, bgg, gbb, gbg, ggb, ggg\}$.

Q.1.5.1.4 Consider the experiment of tossing three coins.
i. Write the sample space.
ii. Find the event of getting exactly one head.
iii. Find the event of getting no head.
Answer: i. Sample space, $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$, where H stands for event of occurring head, and T stands for event of occurring tail
ii. The event of getting exactly one head $= \{HTT, THT, TTH\}$
iii. The event of getting no head $= \{TTT\}$

Q.1.5.1.5 What is the sample space of tossing a coin until one gets an head?
Answer: One might get an head using 1 toss, and the event is $E_1 = \{H\}$.
One might get an head using 2 tosses, and the event is $E_2 = \{TH\}$.
One might get an head using 3 tosses, and the event is $E_3 = \{TTH\}$.
Let E_i be the event of getting head using i tosses.
Thus, the sample space $= \cup_{i=1}^{\infty} E_i = \{H, TH, TTH, \dots\}$

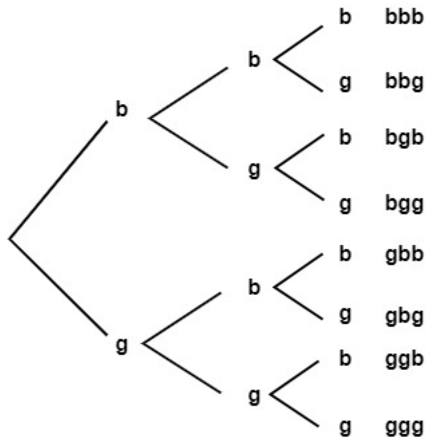

Figure 1.1: Tree diagram of a three-child family

Q.1.5.1.6 Let A and B be events in the same sample space. Then $A \cup (A \cap B) = A$.
Answer: Let Ω be the sample space.
$A \cup (A \cap B) = (A \cap \Omega) \cup (A \cap B) = A \cap (\Omega \cup B)$ [Distributive law, Q.1.5.1.7]
$= A \cap \Omega = A$

Q.1.5.1.7 Consider the events A, B and C in the same sample space. Then $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.
Answer: This is a distributive law. We shall give a prooof of it.
Let $x \in A \cap (B \cup C)$.
Then $x \in A$ and $x \in B \cup C \Rightarrow x \in A$ and ($x \in B$ or $x \in C$)
Case 1: $x \in B$
Then $x \in A$ and $x \in B \Rightarrow x \in A \cap B \Rightarrow x \in (A \cap B) \cup (A \cap C)$
Case 2: $x \in C$
Then $x \in A$ and $x \in C \Rightarrow x \in A \cap C \Rightarrow x \in (A \cap B) \cup (A \cap C)$
Thus, $x \in A \cap (B \cup C) \Rightarrow x \in (A \cap B) \cup (A \cap C)$
Then $A \cap (B \cup C) \subseteq (A \cap B) \cup (A \cap C)$ (1.1)
Let $x \in (A \cap B) \cup (A \cap C)$.
Then $x \in A \cap B$ or $x \in A \cap C$
Case i: $x \in A \cap B$
Then $x \in A$ and $x \in B \Rightarrow x \in A$ and $x \in B \cup C \Rightarrow x \in A \cap (B \cup C)$
Case ii: $x \in A \cap C$

Then $x \in A$ and $x \in C \Rightarrow x \in A$ and $x \in B \cup C \Rightarrow x \in A \cap (B \cup C)$
Then $x \in (A \cap B) \cup (A \cap C) \Rightarrow x \in A \cap (B \cup C)$
Thus, $(A \cap B) \cup (A \cap C) \subseteq A \cap (B \cup C)$ (1.2)
(1.1) and (1.2) $\Rightarrow (A \cap B) \cup (A \cap C) = A \cap (B \cup C)$
Note: Another distributive law: $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

Q.1.5.1.8 Let $A_i = \left[0, 1 - \frac{1}{i}\right]$, for $i = 1, 2, \ldots$.
Find $\cup_{i=1}^{\infty} A_i$ and $\cap_{i=1}^{\infty} A_i$.
Answer: $\cup_{i=1}^{\infty} A_i = [0, 1)$ and $\cap_{i=1}^{\infty} A_i = \{0\}$.

Q.1.5.1.9 If $A_n \subset A_{n-1} \subset \cdots \subset A_1$ show that $\cup_{i=1}^{n} A_i = A_1$ and $\cap_{i=1}^{n} A_i = A_n$.
Answer: We prove the result $\cup_{i=1}^{n} A_i = A_1$ using induction on n.
For $n = 1$, the result is true, by default.
Let the result be true for $n = k$.
Thus, $\cup_{i=1}^{k} A_i = A_1$
We shall show that the result is true for $n = k + 1$.
Then $\cup_{i=1}^{k+1} A_i = \cup_{i=1}^{k} A_i \cup A_{k+1} = A_1 \cup A_{k+1}$ [by induction hypothesis]
$=A_1$ [from the given condition]
Thus, the induction step follows.
[*Second part*]
We prove the result $\cap_{i=1}^{n} A_i = A_n$ using induction on n.
For $n = 1$, the result is true, by default.
Let the result be true for $n = k$.
Thus, $\cap_{i=1}^{k} A_i = A_k$
We shall show that the result is true for $n = k + 1$.
Then $\cap_{i=1}^{k+1} A_i = \cap_{i=1}^{k} A_i \cap A_{k+1} = A_k \cap A_{k+1}$ [by induction hypothesis]
$=A_{k+1}$ [from the given condition]
Thus, the induction step follows.

Q.1.5.1.10 What is a trial? Give an example.
Answer: Consider an experiment, which is being repeated under identical conditions, does not give unique result. But the result is one of the several possible outcomes. The experiment, repeated each time, is known as a trial and the outcome is known as a case or sample point or an elementary event. When we roll a die, an outcome is one of the faces numbered from 1, 2, ..., 6.

Q.1.5.1.11 Explain the following terms: exhaustive events, mutually exclusive events, equally likely events, independent events.

2. Combinatorial Analysis

Q.1.5.2.1 If $A_1, A_2, \ldots, A_k$ are k disjoint sets, then $n(A_1 \cup A_2 \cup \cdots \cup A_k) = n(A_1) + n(A_2) + \cdots + n(A_k)$.
Answer: We shall prove the result using the method of induction on k.
Let $k = 2$.
Event $A_1 \cup A_2$ can occur only by occurring either event A_1 or event A_2, since $A_1 \cap A_2 = \phi$.
Thus, $n(A_1 \cup A_2)$ = The number of sample points in $A_1 \cup A_2$ = The number of sample points in A_1 + The number of sample points in A_2 = $n(A_1) + n(A_2)$.
So, the result is true for $k = 2$.
Let the result be true for $k \leq p$. We want to show that the result is true for $k = p + 1$.
$n(A_1 \cup A_2 \cup \cdots \cup A_{p+1}) = n((A_1 \cup A_2 \cup \cdots \cup A_p) \cup A_{p+1})$
Note that events $A_1 \cup A_2 \cup \cdots \cup A_p$ and A_{p+1} are disjoint.
Then $n(A_1 \cup A_2 \cup \cdots \cup A_{p+1}) = n(A_1 \cup A_2 \cup \cdots \cup A_p) + n(A_{p+1})$ [by induction hypothesis]
$= n(A_1) + n(A_2) + \cdots + n(A_p) + n(A_{p+1})$ [by induction hypothesis]
So, the induction step follows.

Q.1.5.2.2 Prove that $\sum_{k=0}^{n} \binom{N}{k} \binom{M}{n-k} = \binom{N+M}{n}$
for integers $N, M, n \geq 0$.
Answer: [Method 1]
$$(1+x)^{N+M} = (1+x)^N.(1+x)^M = \left(\sum_{k=0}^{N} \binom{N}{k} x^k \right) \left(\sum_{l=0}^{M} \binom{M}{l} x^l \right)$$
Then, $(1+x)^{N+M} = \sum_{k=0}^{N} \sum_{l=0}^{M} \binom{N}{k} \binom{M}{l} x^{k+l}$
We are interested in collecting the coefficients of x^n from both sides for those terms for which $k + l = n$.
$$\sum_{n=0}^{N+M} \binom{N+M}{n} x^n = \sum_{n=0}^{N+M} \sum_{k=0}^{n} \binom{N}{k} \binom{M}{n-k} x^n.$$
The result follows.
[Method 2] Consider a set of $(N + M)$ elements. We divide it into two parts: one with N elements and other with M elements. If we wish to select n elements, then we first choose k elements from the first part

$(k \leq n)$ and the remaining $(n-k)$ elements from the second part. For each k, there are precisely $\begin{pmatrix} N \\ k \end{pmatrix}\begin{pmatrix} M \\ n-k \end{pmatrix}$ possibilities. Summing over all values of k, we obtain an expression for $\begin{pmatrix} N+M \\ n \end{pmatrix}$.

Q.1.5.2.3 State and prove Pascal's identity.
Answer: Pascals's identity states that $\begin{pmatrix} n \\ k \end{pmatrix} = \begin{pmatrix} n-1 \\ k-1 \end{pmatrix} + \begin{pmatrix} n-1 \\ k \end{pmatrix}$, for any positive integers k and n.
[Method 1]
When $k > n$, $\begin{pmatrix} n \\ k \end{pmatrix} = 0 = \begin{pmatrix} n-1 \\ k-1 \end{pmatrix} + \begin{pmatrix} n-1 \\ k \end{pmatrix}$
Thus, the result is true.
Now we assume that $k \leq n$.
$\begin{pmatrix} n-1 \\ k-1 \end{pmatrix} + \begin{pmatrix} n-1 \\ k \end{pmatrix} = \frac{(n-1)!}{(k-1)!(n-k)!} + \frac{(n-1)!}{k!(n-k-1)!}$
$= (n-1)! \left[\frac{k}{k!(n-k)!} + \frac{n-k}{k!(n-k)!}\right] = (n-1)! \left[\frac{n}{k!(n-k)!}\right] = \frac{n!}{k!(n-k)!} = \begin{pmatrix} n \\ k \end{pmatrix}$
[Method 2]
For $k = 0$, $\begin{pmatrix} n-1 \\ k-1 \end{pmatrix} = \begin{pmatrix} n-1 \\ -1 \end{pmatrix} = 0$, and $\begin{pmatrix} n \\ 0 \end{pmatrix} = \begin{pmatrix} n-1 \\ 0 \end{pmatrix}$.
For $k = n$, $\begin{pmatrix} n-1 \\ k \end{pmatrix} = \begin{pmatrix} n-1 \\ n \end{pmatrix} = 0$, and $\begin{pmatrix} n \\ n \end{pmatrix} = \begin{pmatrix} n-1 \\ n-1 \end{pmatrix}$.
The proof of the above result is give below for $0 < k < n$.
Let $\{a_1, a_2, \ldots, a_n\}$ be the set of elements under consideration. There are two types of k-subsets: subsets that contain a_n and subsets that do not contain a_n.
Case 1: When a k-subset does not contain a_n, then the k-subset is formed from $\{a_1, a_2, \ldots, a_{n-1}\}$. The number of such subsets is $\begin{pmatrix} n-1 \\ k \end{pmatrix}$.
Case 2: When a k-subset contains a_n, then $k-1$ elements are selected from $\{a_1, a_2, \ldots, a_{n-1}\}$. The number of such subsets is $\begin{pmatrix} n-1 \\ k-1 \end{pmatrix}$.
Adding the k-subsets of Case 1 and Case 2, we get the total number of k-subsets. The result follows.

Q.1.5.2.4 State binomial theorem. Prove it.
Answer: Binomial theorem states that $(x+y)^n = \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} x^k y^{n-k}$, for $n = 1, 2, 3, \ldots$ (2.1)

[Method 1]
For $n = 1$, LHS $= x + y$.
RHS $= \begin{pmatrix} 1 \\ 0 \end{pmatrix} x^0 y^{1-0} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} x^1 y^{1-1} = y + x =$ LHS
The theorem holds for $n = 1$. Assume that the theorem holds for $n - 1$. We need to show that the theorem holds for n.
$(x+y)^n = (x+y)(x+y)^{n-1} = (x+y) \sum_{k=0}^{n-1} \begin{pmatrix} n-1 \\ k \end{pmatrix} x^k y^{n-1-k}$
$= \sum_{k=0}^{n-1} \begin{pmatrix} n-1 \\ k \end{pmatrix} x^{k+1} y^{n-1-k} + \sum_{k=0}^{n-1} \begin{pmatrix} n-1 \\ k \end{pmatrix} x^k y^{n-k}$
Let $i = k + 1$ for the first sum and $i = k$ for the second sum.
$(x+y)^n = \sum_{i=1}^{n-1} \begin{pmatrix} n-1 \\ i-1 \end{pmatrix} x^i y^{n-i} + \sum_{i=0}^{n-1} \begin{pmatrix} n-1 \\ i \end{pmatrix} x^i y^{n-i}$
$= x^n + \sum_{i=1}^{n-1} \left[\begin{pmatrix} n-1 \\ i-1 \end{pmatrix} + \begin{pmatrix} n-1 \\ i \end{pmatrix} \right] x^i y^{n-i} + y^n$
$= x^n + \sum_{i=1}^{n-1} \begin{pmatrix} n \\ i \end{pmatrix} x^i y^{n-i} + y^n$ [Pascal's identity, Q.1.5.2.3]
$= \sum_{i=0}^{n} \begin{pmatrix} n \\ i \end{pmatrix} x^i y^{n-i}$
[Method 2]
In the expansion of $(x_1 + y_1)(x_2 + y_2) \dots (x_n + y_n)$, each term contains n of x_i and y_i, $i = 1, 2, \dots, n$. There are 2^n such terms. When $n = 2$, we have $(x_1 + y_1)(x_2 + y_2) = x_1y_1 + x_1y_2 + x_2y_1 + x_2y_2$.
When $x_1 = x_2 = \dots = x_n = x$ and $y_1 = y_2 = \dots = y_n = y$, how many of the 2^n terms in the sum will have k of the xs and $(n - k)$ of the ys as factors? The form of such term is $x^k y^{n-k}$, and the number of such terms is equal to the number of choices of finding k x's out of possible n x's, i.e., $\begin{pmatrix} n \\ x \end{pmatrix}$.

Thus, $(x+y)^n = \sum_{k=0}^{n} \begin{pmatrix} n \\ k \end{pmatrix} x^k y^{n-k}$.

Q.1.5.2.5 How many integers between 1 and 300 (inclusive) are
i. divisible by at least one of 3, 5, 7?
ii. divisible by 3 and by 5 but not by 7?
iii. divisible by 5 but by neither 3 nor 7?
Answer: Let $A = \{n | 1 \leq n \leq 300 \wedge 3 \textit{ divides } n\}$,

3. Probability Models

Q.1.5.3.1 What is a probability model?
Answer: A probability model consists of a nonempty set called the sample space S; a collection of events that are subsets of S; and a probability measure P having the following properties:
i. $P(A) \in [0,1]$, where event A is defined on the sample space S.
ii. $P(S) = 1$.
iii. P is countably additive. Let $A_1, A_2, \dots$ be a finite or countable sequence of disjoint events.
Then $P(A_1 \cup A_2 \cup \dots) = P(A_1) + P(A_2) + \dots$ (3.1)
(i) - (iii) are called axioms of probability.

Q.1.5.3.2 Show that $P(\emptyset) = 0$, where $\emptyset$ is null set.
Answer: Let $E_1, E_2, E_3, \dots$ be the sequence of events such that $S = E_1$, and $E_2 = E_3 = \dots = \emptyset$, where S is the sample space.
Events are mutually exclusive, and $\bigcup_{i=1}^{\infty} E_i = S$.
Using axiom (iii), we have
$P(S) = P(E_1) + P(E_2) + P(E_3) + \dots = P(S) + P(E_2) + P(E_3) + \dots$
$\Rightarrow P(S) = P(S) + P(\emptyset) + P(\emptyset) + \dots$
$\Rightarrow P(\emptyset) = 0$, where $\emptyset \subseteq S$, and $P(\emptyset) \geq 0$ [Axiom (i)]

Q.1.5.3.3 Let E be an event defined on sample space S. Show that $P(E^c) = 1 - P(E)$, where $P(E^c)$ is the complement event of E.
Answer: $1 = P(S)$ [Axiom (ii)]
$\Rightarrow 1 = P(E \cup E^c)$ [$S = E \cup E^c$]
$\Rightarrow 1 = P(E) + P(E^c)$ [Axiom (iii)]
$\Rightarrow P(E^c) = 1 - P(E)$

Q.1.5.3.4 Show that $P(E) \leq P(F)$, if $E \subseteq F$.
Answer: $F = E \cup (\bar{E} \cap F)$, since $E \subseteq F$
Note that events E and $\bar{E} \cap F$ are mutually exclusive.
Using Axiom (iii), $P(F) = P(E) + P(E^c \cap F)$ (3.2)
Using Axiom (i), $P(E^c \cap F) \geq 0$.
From (3.2), $P(F) \geq P(E)$.

Q.1.5.3.5 What is relative frequency? Give an example.
Answer: Consider an experiment that is performed n times. Each per-

formance is called a trial (see Q.1.5.1.10). Let an event E occur in n_E trials. Then the ratio $\frac{n_E}{n}$ is called the relative frequency of E in the n trials, and denoted by $f(E, n) = \frac{n_E}{n}$.
An unbiased coin is tossed 100 times, and tail (T) appears 52 times. Then the relative frequency of T is $52/100 = 0.52$.
When n is a large number, $P(E) \approx f(E, n)$.

Q.1.5.3.6 State and prove the formula of probability for union of events.
Answer: *Probability of union of events*: For any n events $E_1, E_2, \dots, E_n$, $P(E_1 \cup E_2 \cup \dots \cup E_n) = \sum_i P(E_i) - \sum_{i<j} P(E_i \cap E_j) + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) - \dots + (-1)^{n+1} P(E_1 \cap E_2 \cap \dots \cap E_n)$
Proof: We shall prove it using the method of induction on n. Let $n = 2$.
Then $P(E_1 \cup E_2) = P[(E_1 \cap E_2^c) \cup (E_1^c \cap E_2) \cup (E_1 \cap E_2)]$
Events $E_1 \cap E_2^c$, $E_1^c \cap E_2$ and $E_1 \cap E_2$ are mutually exclusive.
Thus, $P(E_1 \cup E_2) = P(E_1 \cap E_2^c) + P(E_1^c \cap E_2) + P(E_1 \cap E_2)$, (3.3)
using Axiom (iii), Q.1.5.3.1
$P(E_1) = P[(E_1 \cap E_2) \cup (E_1 \cap E_2^c)]$
$= P(E_1 \cap E_2) + P(E_1 \cap E_2^c)$ [Axiom (iii)]
Then $P(E_1 \cap E_2^c) = P(E_1) - P(E_1 \cap E_2)$
Similarly, $P(E_1^c \cap E_2) = P(E_2) - P(E_1 \cap E_2)$
From (3.3),
$P(E_1 \cup E_2) = P(E_1) - P(E_1 \cap E_2) + P(E_2) - P(E_1 \cap E_2) + P(E_1 \cap E_2)$
So, $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$
Thus, the result is true for $n = 2$.
We assume that the result is true for $n \leq k$. We shall show that the result is true for $n = k + 1$.
$P(E_1 \cup E_2 \cup \dots \cup E_k \cup E_{k+1}) = P[(E_1 \cup E_2 \cup \dots \cup E_k) \cup E_{k+1}]$
$= P(E_1 \cup E_2 \cup \dots \cup E_k) + P(E_{k+1}) - P[(E_1 \cup E_2 \cup \dots \cup E_k) \cap E_{k+1}]$,
using induction hypothesis
$= \sum_i P(E_i) - \sum_{i<j} P(E_i \cup E_j) + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) - \dots + (-1)^{n+1} P(E_1 \cap E_2 \cap \dots \cap E_n) + P(E_{k+1}) - T$, (3.4)
where, $T = P[(E_1 \cap E_{k+1}) \cup (E_2 \cap E_{k+1}) \cup \dots \cup (E_k \cap E_{k+1})]$,
applying distributive law
$T = \sum_i P(E_i \cap E_{k+1}) - \sum_{i<j} P(E_i \cap E_j \cap E_{k+1}) + \dots$, (3.5)
applying induction hypothesis
Using (3.4) and (3.5), we conclude that the induction step follows.

Q.1.5.3.7 Discuss classical (apriori) concept of probability. State the limitations of the concept.
Answer: Assume that a trial in an experiment results in one of n ex-

haustive, mutually exclusive and equally likely cases, and m of them are favourable to the happening of an event E.

$P(E) = \frac{Number\ of favourable\ cases}{Number\ of\ exhaustive\ cases} = \frac{m}{n}$

There are limitations of the above definition.
i. Outcomes of a trial may not be equally likely.
ii. The number of exhaustive cases in a trial may be infinite.

Q.1.5.3.8 Elaborate the emperical (statistical) concept of probability.
Answer: Consider an experiment, where the trials are repeated under essentially homogeneous and identical conditions. Then the limiting value of the ratio of the number of times an event happens to the number of trials, when the number of trials become indefinitely large, is called the probability of the the event.
Refer to the concept of relative frequency (Q.1.5.3.5). Let an event E occurs n_E times in n trials. Let $f(E, n)$ be the relative frequency of event E in n trials. Then emperical probability of E is defined as
$P(E) = \lim_{n\to\infty} f(E, n)$.

Q.1.5.3.9 Let $E_1, E_2, \ldots, E_n$ be n events on the same sample space. Prove that $P\left(\bigcup_{i=1}^{n} E_i\right) \leq \sum_{i=1}^{n} P(E_i)$.
Answer: We shall prove it using the method of induction on n.
When $n = 2$, $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$ [Q.1.5.3.6]
Using Axiom (i), $P(E_1 \cap E_2) \geq 0$ [Q.1.5.3.1]
Then, $P(E_1 \cup E_2) \leq P(E_1) + P(E_2)$
We assume that the result is true for $n = k$. We shall prove that the result is true for $n = k + 1$.
$P(E_1 \cup E_2 \cup \cdots \cup E_k \cup E_{k+1}) = P[(E_1 \cup E_2 \cup \cdots \cup E_k) \cup E_{k+1}]$
$= P(E_1 \cup E_2 \cup \cdots \cup E_k) + P(E_{k+1}) - P[(E_1 \cup E_2 \cup \cdots \cup E_k) \cap E_{k+1}]$ [Q.1.5.3.6]
$\leq \sum_{i=1}^{k} P(E_i) + P(E_{k+1}) - P[(E_1 \cup E_2 \cup \cdots \cup E_k) \cap E_{k+1}]$ [by induction hypothesis]
$\leq \sum_{i=1}^{k+1} P(E_i)$ [Axiom (i), Q.1.5.3.1]

Q.1.5.3.10 Three unbiased coins are flipped. Find the probability that first and third coins show the same face.
Answer: Sample space $= \{(H, H, H), (H, H, T), (H, T, H), (H, T, T), (T, H, H), (T, H, T), (T, T, H), (T, T, T)\}$
Let E be the event that first and the third coins turn into the same face.