MT241P FINITE MATHEMATICS

1. Natural Numbers, Integers and Rational Numbers

Definition 1.1. *A* **Ring** *is a set R together with two operations of addition $+$ and multiplication $\cdot$, where the following rules hold, for $a, b, c \in R$:*

(R1) $a + b = b + a$ *(addition is commutative);*
(R2) $a + (b + c) = (a + b) + c$ *(addition is associative);*
(R3) $a + 0 = a = 0 + a$ *(0 is the additive identity);*
(R4) *there exists an integer x such that $a + x = 0 = x + a$. We denote x by $-a$ (existence of additive inverses);*
(R5) $a \cdot b = b \cdot a$ *(multiplication is commutative);*
(R6) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ *(multiplication is associative);*
(R7) $a \cdot 1 = a = 1 \cdot a$ *(1 is the multiplicative identity);*
(R8) $a \cdot (b + c) = a \cdot b + a \cdot c$ *and* $(a + b) \cdot c = a \cdot c + b \cdot c$ *(multiplication distributes over addition);*

Remark 1.2. (Natural Numbers)

In the following we want to define the natural numbers using the five **Peano Axioms**. *To this end let X be a set such that :*

(PA1) $0 \in X$, *that is, X has at least one element*
(PA2) *there is a function $S : X \to X$, called the* **successor function**,
(PA3) *there is no $x \in X$ such that $S(x) = 0$,*
(PA4) *if $S(x) = S(y)$, then $x = y$,*
(PA5) *(Axiom of Induction) If K is a subset of X such that*
(a) $0 \in K$ and
(b) if $x \in K$ then $S(x) \in K$,
then $K = X$.

By (PA1) there is an element $0 \in X$. (PA2) implies that $S(0)$ exists and it differs from 0, by (PA3). Now, by (PA2), an element $S(S(0))$ exists, which differs from 0, by (PA3), and differs from $S(0)$, by (PA4). Going on like this we get a sequence of distinct elements

$$(\star) \quad 0, S(0), S(S(0)), S(S(S(0))), \ldots$$

This alone will not describe the natural numbers as we know them, because so far all of this also holds true for instance for the interval $[0, \infty)$ of real *numbers, where* 0 *is the real number zero and* $S(x) = x + 1$*. However, (PA5) ensures that every element in* X *is the successor of another element. Because let* K *be the list* $(\star)$ *of elements. Then* K *satisfies conditions (PA5a) and (PA5b), and therefore* $K = X$*. In particular* X *is restricted to only elements on the list and therefore cannot be* $[0, \infty)$.

We agree to the following notation:

$$1 := S(0),\ 2 = S(1),\ 3 = S(2), \ldots$$

and we denote X *by* $\mathbb{N}$ *and call it the* **natural numbers***. In particular* $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$.

We define an addition (+) and a multiplication (·) on $\mathbb{N}$*. For all* $m, n \in \mathbb{N}$*:*

$$(A1)\ m + 0 := m \qquad (M1)\ m \cdot 0 := 0$$

$$(A2)\ m + S(n) := S(m + n) \qquad (M2)\ m \cdot S(n) := m + (m \cdot n)$$

One can show that $(\mathbb{N}, +, \cdot)$ *satisfies all ring properties except for (R4).*

Remark 1.3. (Integers)

For two $m, n \in \mathbb{N}$ *the equation* $m + x = n$ *may or may not have a solution for* x *in* $\mathbb{N}$*. Let the pair* $(n, m) \in \mathbb{N} \times \mathbb{N}$ *represent the equation. Then we say two equations* $(n_1, m_1), (n_2, m_2)$ *are* **equivalent** *if* $n_1 + m_2 = n_2 + m_1$*. One can show that the elements*

$$X := \{(0, 0), (1, 0), (0, 1), (2, 0), (0, 2), (3, 0), (0, 3), \ldots\}$$

are all non-equivalent, but every element $(n, m) \in \mathbb{N} \times \mathbb{N}$ *is equivalent to precisely one of them. Furthermore we define an addition and multiplication on* $\mathbb{N} \times \mathbb{N}$ *as follows:*

(Addition) $(n_1, m_1) + (n_2, m_2) := (n_1 + n_2, m_1 + m_2)$

(Multiplication) $(n_1, m_1) \cdot (n_2, m_2) := (n_1 \cdot n_2 + m_1 \cdot m_2, n_1 \cdot m_2 + n_2 \cdot m_1)$, *for all* $n_1, n_2, m_1, m_2 \in \mathbb{N}$.

If we identify each (n, m) *with its equivalence companion in* X*, then one can show that* $(X, +, \cdot)$ *is a ring. If we identify the element* $(n, 0)$ *with* $n \in \mathbb{N}$*, then* X *contains* $\mathbb{N}$*. Furthermore* $(0, n)$ *is the additive inverse of* $(n, 0)$ *and thus can be written as* $-n$*. Overall we denote the set* X *by* $\mathbb{Z}$ *and call it the* **integers***. In particular we have*

$$\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$

There is an **order** *on* $\mathbb{Z}$*. For* $a, b \in \mathbb{Z}$ *we write* $a < b$ *if there is an* $x \in \mathbb{N}\backslash\{0\}$ *such that* $a + x = b$*, or in other words* $(-a) + b \in \mathbb{N} \setminus \{0\}$.

If $(-a) + b \notin \mathbb{N} \setminus \{0\}$*, then either* $(-a) + b = 0$*, that is,* $a = b$*, or* $(-a) + b = -c$*, for some* $c \in \mathbb{N} \setminus \{0\}$*. The latter gives* $b + c = a$ *and therefore* $b < a$*. Thus* *precisely one of the following is true:*

$$a < b, \quad a = b, \quad b < a.$$

Furthermore we write $a \leq b$ *if* $a < b$ *or* $a = b$*, we write* $a > b$*, if* $b < a$ *and we write* $a \geq b$ *if* $b \leq a$*.*

Next note that for every $a \in \mathbb{Z}$ *we have* $0 < a$ *if and only if* $a \in \mathbb{N} \setminus \{0\}$ *and we call those integers* **positive***. Moreover* $a < 0$ *if and only* $a \in \mathbb{Z} \setminus \mathbb{N}$ *and we call those integers* **negative***.*

Remark 1.4. (Rational Numbers) *Observe that the equation* $3 \cdot x = 4$ *has no solution in* $\mathbb{Z}$*. We identify the equation* $b \cdot x = a$*, where* $a, b \in \mathbb{Z}, b \neq 0$ *with the pair* (a, b) *and we say two such pairs* (a, b) *and* (c, d) *are equivalent if* $a \cdot d = b \cdot c$*. Set*

$$X := \{(a, b) : a, b \in \mathbb{Z}, b \neq 0\},$$

where we understand equivalent elements as the same element. Then X *forms a ring under the operations*

$$(a, b) + (c, d) = (ad + bc, bd), \qquad (a, b) \cdot (c, d) = (ac, bd).$$

The additive identity is $(0, 1)$ *(which equals* $(0, x)$*, for any integer* $x \neq 0$*) and the multiplicative identity is* $(1, 1)$ *(which equals* (x, x)*, for any* $x \neq 0$*). Overall we denote the set* X *by* $\mathbb{Q}$ *and call it the* **rational numbers***. Furthermore we denote the pair* (a, b)*, by the fraction* $\frac{a}{b}$*. Overall we get that*

$$\mathbb{Q} = \left\{ \frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0 \right\}.$$

In addition, the rationals are ordered by

$$\frac{a}{b} < \frac{c}{d} \quad \text{if and only if} \quad a \cdot d < b \cdot c.$$

We can identify $\mathbb{Z}$ *as the subset* $\{\frac{a}{1} \mid a \in \mathbb{Z}\}$ *of* $\mathbb{Q}$*. Note that then the above operations and order are an extension of the operations and order on* $\mathbb{Z}$*.*

Note that by (A1) and (A2) we get for every $m \in \mathbb{N}$ that

$$S(m) = S(m+0) = m + S(0) = m + 1.$$

Hence the Axiom of Induction can be rephrased as follows:

(PA5) If K is a subset of $\mathbb{N}$ such that
(a) $0 \in K$ and
(b) if $n \in K$ then $n+1 \in K$,
then $K = \mathbb{N}$.

Theorem 2.1. (Principle of Induction)
Let $n_0 \in \mathbb{Z}$ and let $P(n)$ be a property which can be true or false, for all $n \in \mathbb{Z}$ with $n_0 \leq n$. Furthermore assume that $P(n_0)$ is true and if $P(n)$ is true for some $n \in \mathbb{Z}$ with $n_0 \leq n$, then $P(n+1)$ is true. Then $P(n)$ is true, for all integers n with $n_0 \leq n$.

Proof. Set $K := \{n \in \mathbb{N} : P(n_0 + n) \text{ is true.}\}$. As $P(n_0)$ is true we have $0 \in K$. Furthermore if $n \in K$, then $P(n_0 + n)$ is true and so by assumption $P(n_0 + n + 1)$ is true. Consequently, by the axiom of induction, $n + 1 \in K$ and thus $K = \mathbb{N}$. □

Example 2.2. *(1) Show that $\sum_{i=1}^{n} i = 1+2+\ldots+n = \frac{n(n+1)}{2}$, for all integers $n \geq 1$. The statement is true for $n = 1$, as the sum on the left and the fraction on the right both equal one. Next suppose the statement is true for some $n \geq 1$. We need to show it holds for $n+1$. We have*

$$\begin{aligned}\sum_{i=1}^{n+1} i &= (n+1) + \sum_{i=1}^{n} i = (n+1) + \frac{n(n+1)}{2} = (n+1)\cdot\left(1 + \frac{n}{2}\right) \\ &= (n+1)\cdot\left(\frac{n+2}{2}\right) = \frac{(n+1)(n+2)}{2}.\end{aligned}$$

Hence the statement is true for $n+1$. In particular, the statement is true for all $n \geq 1$.

(2) For $n, k \in \mathbb{N}$ such that $0 \leq k \leq n$ we define the **binomial coefficient**

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

We claim that all binomial coefficient are integers. This statement is quickly verified to be true if $n = 0$ or $k = 0$. Henceforth we assume that neither are zero. If $n = 1$, then $k = 1$ and as $\binom{1}{1} = 1$. Hence the statement holds for

$n = 1$. Next assume the statement is true for some integer $n \geq 1$ and all integers k with $0 \leq k \leq n$. Then for all $1 \leq k \leq n+1$ we have

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

Note that both $\binom{n}{k}$ and $\binom{n}{k-1}$ are integers and consequently so is $\binom{(n+1)}{k}$.

Corollary 2.3. (Strong Induction)

Let $n_0 \in \mathbb{Z}$ and let $P(n)$ be a property which can be true or false, for all $n \in \mathbb{Z}$ with $n_0 \leq n$. Furthermore assume that $P(n_0)$ is true and that $P(n+1)$ is true, for some $n \in \mathbb{Z}$, whenever $P(k)$ is true for all integers k with $n_0 \leq k \leq n$. Then $P(n)$ is true, for all integers n with $n_0 \leq n$.

Example 2.4. *Set $T_1 = 1, T_2 = 3$ and recursively $T_n = T_{n-1} + T_{n-2}$, for $n \geq 3$ (Lucas Sequence). Prove that*

$$T_n < \left(\frac{7}{4}\right)^n, \quad \textit{for all } n \geq 1.$$

As

$$T_1 = 1 < \frac{7}{4} \textit{ and } T_2 = 3 < \frac{49}{16} = \left(\frac{7}{4}\right)^2,$$

the statement holds for 1 and 2. Next suppose it holds for all integers k where $1 \leq k \leq n$, for some $n \geq 2$. Then

$$T_{n+1} = T_n + T_{n-1} < \left(\frac{7}{4}\right)^n + \left(\frac{7}{4}\right)^{n-1} = \left(\frac{7}{4}\right)^{n-1} \cdot \left(\frac{7}{4} + 1\right)$$
$$= \left(\frac{7}{4}\right)^{n-1} \cdot \left(\frac{11}{4}\right) = \left(\frac{7}{4}\right)^{n-1} \cdot \left(\frac{44}{16}\right) < \left(\frac{7}{4}\right)^{n-1} \cdot \left(\frac{49}{16}\right)$$
$$= \left(\frac{7}{4}\right)^{n-1} \cdot \left(\frac{7}{4}\right)^2 = \left(\frac{7}{4}\right)^{n+1}$$

Hence the statement holds for $n+1$ and thus for all $n \geq 1$.

Theorem 2.5. (*Well-Ordering Principle) Let S be a non-empty subset of $\mathbb{N}$. Then S has a least element, that is, there is some $l \in S$ so that $l \leq s$, for all $s \in S$.*

Proof: Let S be a subset of $\mathbb{N}$ without a least element. Then surely $0 \notin S$. Now set

$$K := \{n \in \mathbb{N} : n \notin S\}$$

We say $P(n)$ is true for some $n \in \mathbb{N}$, if $n \in K$. Hence $P(0)$ is true. Next let $n \geq 0$ and assume that $P(k)$ is true for all integers k with $0 \leq k \leq n$. If $n+1 \in S$, then $n+1$ will be a least element of S, as otherwise there must be an $s \in S$ with $s < n+1$. But then $P(s)$ is true by assumption, that is, $s \in K$ in contradiction to $s \in S$. Therefore $n+1 \notin S$, that is, $n+1 \in K$. Now by Corollary 2.3 we get that $K = \mathbb{N}$. In particular, $S = \emptyset$.

□

Corollary 2.6. *There is no $a \in \mathbb{Z}$ such that $0 < a < 1$.*

Proof: Assume there is such an $a \in \mathbb{Z}$ and let S be the set of all such a. Then S is a non-empty subset of $\mathbb{N}$ and as such contains a least element l. Note that $l \in \mathbb{N} \setminus \{0\}$. As $l < 1$, there is $x \in \mathbb{N} \setminus \{0\}$ such that $l + x = 1$. Multiplying l onto this equation gives $l^2 + xl = l$. Clearly $xl \in \mathbb{N} \setminus \{0\}$ and so $l^2 < l$. We also have $0 < l^2 < 1$, that is $l^2 \in S$. But this contradicts the minimality of l. In particular there is no a as described.

□

Remark 2.7. *The integers do not satisfy the Well-Ordering Principle, as there are subsets S of $\mathbb{Z}$ that do not contain a smallest element, take for instances $S = \mathbb{Z}$ or the subset of negative integers. In particular any number system containing the integers, such as $\mathbb{Q}$ for instance, cannot satisfy the Well-Ordering Principle either.*

3. The Division Algorithm

Lemma 3.1. *Let $a, b, c \in \mathbb{Z}$ such that $a < b$. Then*

1. $a + c < b + c$.
2. *If* $0 < c$, *then* $ac < bc$.
3. *If* $c < 0$, *then* $bc < ac$.

Proof: By assumption $a + x = b$, for some $x \in \mathbb{N} \setminus \{0\}$.
(1) Then $(a + c) + x = (b + c)$ and so $a + c < b + c$ ensues.

(2) We have $(a + x)c = bc$ and so $ac + xc = bc$. Since $c > 0$ we have $c \in \mathbb{N} \setminus \{0\}$ and so $xc \in \mathbb{N} \setminus \{0\}$. Thus $ac < bc$.

(3) By (2) we have $a(-c) < b(-c)$. Adding $ac+bc$ onto both sides, it follows from part (1), that $bc < ac$.

□

Definition 3.2. *Let $a, b \in \mathbb{Z}$. We say a* **divides** *b and write $a \mid b$, if there is some $c \in \mathbb{Z}$ such that $ac = b$. Alternatively, we say that a is a* **factor** *of b and b is a* **multiple** *of a.*

Lemma 3.3. *Let $a, b, c \in \mathbb{Z}$. Then*

1. $a \mid 0$
2. *If* $0 \mid a$, *then* $a = 0$
3. $a \mid a$
4. $1 \mid a$
5. *If* $a \mid b$ *and* $b \neq 0$, *then* $a \leq |b|$.
6. *If* $a \mid b$ *and* $b \mid c$, *then* $a \mid c$.
7. *If* $a \mid b$ *and* $a \mid c$ *and* $u, v \in \mathbb{Z}$, *then* $a \mid (ub + vc)$.

Proof: (1) $a \cdot 0 = 0$
(2) By assumption $0 \cdot c = a$, for some $c \in \mathbb{Z}$. Hence $a = 0 \cdot c = 0$.
(3) $a \cdot 1 = a$
(4) $1 \cdot a = a$
(5) As surely $a \leq |a|$ it suffices to show that $|a| \leq |b|$. By assumption $ax = b$, for some $x \in \mathbb{Z}$ and so $|a|x = |b|$, for some $x \in \mathbb{N}$. Note that $x \neq 0$, as otherwise $b = 0$. In particular $x \geq 1$. If $x = 1$, then $|a| = |b|$. If $x > 1$, then $|b| = |a|(1 + x - 1) = |a| + |a|(x - 1)$. Note that $|a|(x - 1) \in \mathbb{N} \setminus \{0\}$, and so $|a| < |b|$. Overall $|a| \leq |b|$, as required.
(6) + (7) Homework.

Example 3.4. *We use induction to show that the value* $23^n - 1$ *is divisible by* 11*, for all* $n \geq 0$*. For* $n = 0$ *we get that* $23^n - 1 = 1 - 1 = 0$*, which is divisible by* 11*. Next assume that* $n \geq 0$ *such that* $11 \mid 23^n - 1$*. Then*

$$23^{n+1} - 1 = 23 \cdot 23^n - (23 - 22) = 23 \cdot (23^n - 1) + 22$$

By assumption $11 \mid 23^n - 1$ *and, since* $22 = 11 \cdot 2$ *we have* $11 \mid 22$*. Now by Lemma 3.3(7) we get that* $11 \mid 23 \cdot (23^n - 1) + 22 = 23^{n+1} - 1$*.*

Theorem 3.5 (Division Algorithm)**.** *Let* $a, b \in \mathbb{Z}$ *with* $b \neq 0$*. Then there exist unique* $q, r \in \mathbb{Z}$ *such that*

$$a = qb + r, \quad \text{and } 0 \leq r < |b|.$$

Proof: W.l.o.g $b > 0$. Let $S = \{s \in \mathbb{N} : s = a - qb, \text{ for some } q \in \mathbb{Z}\}$. We claim that S is non-empty. If $a \geq 0$, then choose $q = 0$ and so $s = a - qb = a \in \mathbb{N}$. If $a < 0$, we choose $q = a$. Then $s = a - qb = a - ab = (-a)(-1) + (-a)b = (-a)(b - 1)$. But $1 \leq b$, by Corollary 2.6, and so $b - 1 \geq 0$. Since $-a > 0$, we get $s \in \mathbb{N}$.

Now S contains a minimal element r, by Theorem 2.5. Note that $r \geq 0$ and $r = a - qb$, for some $q \in \mathbb{Z}$, or in other words $a = qb + r$. Next we show that $r < b$. Assume otherwise, that is, $b \leq r$. Then $0 \leq r - b$ and so

$$0 \leq r - b = (a - qb) - b = a - b(q + 1).$$

But then $r - b \in S$ and surely $r - b < r$, in contradiction to the minimality of r. In particular we must have $r < b$.

It remains to show uniqueness. So let's have $a = q'b + r'$, for $q', r' \in \mathbb{Z}$ and $0 \leq r' < b$. Without lose of generality we assume that $r \leq r'$. Then

$$0 = a - a = qb + r - (q'b + r') = (q - q')b + (r - r'),$$

and so

$$0 \leq r' - r = (q - q')b.$$

As $b > 0$ we must have that $(q - q') \geq 0$, by Lemma 3.1. On the other hand $r' < b$ and so $r' - r < b - r < b$, by Lemma 3.1. Therefore $(q - q')b < b$ and so $q - q' = 0$. Hence $q = q'$ and consequently $r = r'$.

□

Example 3.6. *(1) We have* $15 = 2 \cdot 6 + 3$*, is the only way to write* 15 *as a multiple of* 6 *plus a natural number less than* 6*.*

(2) Let $n \in \mathbb{Z}$*. Then there are* $q \in \mathbb{Z}$ *and* $r \in \{0, 1\}$ *such that* $n = 2q + r$*. In other words either* $n = 2q$*, in which case we call* n **even***, or* $n = 2q + 1$*, in*

which case we call n **odd***.*

(3) We show that every odd square is of the form $8n+1$, for some integer n. Let $k = 2q+1$ be odd. Then $k^2 = (2q+1)^2 = 4q^2+4q+1 = 4q(q+1)+1$. Note that either q or $q+1$ is even and therefore $q(q+1) = 2n$, for some $n \in \mathbb{Z}$. Finally we get that $k^2 = 4q(q+1)+1 = 4 \cdot 2n+1 = 8n+1$.

(4) Let $n \in \mathbb{Z}$. Then $n(n^2+2)/3 \in \mathbb{Q}$. But is it in $\mathbb{Z}$? There are $q \in \mathbb{Z}$ and $r \in \{0,1,2\}$ such that $n = 3q+r$, that is, $n = 3q$ or $n = 3q+1$ or $n = 3q+2$. If $n = 3q$, then surely $3 \mid n$. If $n = 3q+1$, then $n^2+2 = (3q+1)^2+2 = 2 = 9q^2+6q+3 = 3(3q^2+2q+1)$ and so $3 \mid n^2+2$. Finally if $n = 3q+2$, then $n^2+2 = (3q+2)^2+2 = 2 = 9q^2+12q+6 = 3(3q^2+4q+2)$ and so again $3 \mid n^2+2$. Overall we conclude that $3 \mid n(n^2+2)$, that is $n(n^2+2)/3$ is an integer, for all $n \in \mathbb{Z}$.

4. Greatest Common Divisor

Definition 4.1. *For $a, b \in \mathbb{Z}$, not both zero, we call $d \in \mathbb{Z}$ their* **greatest common divisor** *and write $d = \gcd(a, b)$ if:*

1. $d > 0$;
2. $d \mid a$ *and* $d \mid b$;
3. *if $c \in \mathbb{Z}$ is such that $c \mid a$ and $c \mid b$, then $c \leq d$.*

We say that a, b are **coprime** *if* $\gcd(a, b) = 1$.

Lemma 4.2. *Let $a, b \in \mathbb{Z}$, such that $a \neq 0$ and $a \mid b$. Then $\gcd(a, b) = |a|$.*

Proof: Clearly $|a| > 0$ and $|a|$ divides both a and b. Furthermore if $c \in \mathbb{Z}$ is such that $c \mid a$ and $c \mid b$, then by Lemma 3.3(5), we get that $c \leq |a|$. In particular $\gcd(a, b) = |a|$.

□

Theorem 4.3. *For $a, b \in \mathbb{Z}$, not both zero, $\gcd(a, b)$ exists and is the minimal element of $S := \{sa + tb \mid s, t \in \mathbb{Z}, sa + tb > 0\}$.*

Proof: Note that $|a|, |b| \in S$ and so S is a non-empty set of natural numbers. Then S has a minimal element $d = sa + tb$. Clearly $d > 0$. Next we show that $d \mid a$. Assume it does not. By the Division Algorithm there are $q, r \in \mathbb{Z}$ such that $a = qd + r$, with $0 \leq r < d$. As $d \nmid a$ we must have $0 < r$. Now

$$r = a - qd = a - q(sa + tb) = (1 - qs)a + (-qt)b.$$

But then $r \in S$ and $r < d$, which contradicts the minimality of d. Hence $d \mid a$. Likewise one shows that $d \mid b$.

Finally let $c \in \mathbb{Z}$ such that $c \mid a$ and $c \mid b$. If $c \leq 0$, then surely $c < d$. Hence assume that $c > 0$. By Lemma 3.3 (7) we get $c \mid d$, that is, $ct = d$, for some $t \in \mathbb{Z}$. By Lemma 3.1 we have $t > 0$, as otherwise $d < 0$. Hence $d = c(1 + (t - 1)) = c + c(t - 1)$. Now either $t - 1 = 0$, in which case $d = c$ or $t - 1 > 0$, in which case $c < d$. In all cases we have $c \leq d$. Over all this shows that $d = \gcd(a, b)$.

□

Corollary 4.4. *Let a, b be integers, not both zero. Then*

1. *a and b are coprime if and only there exist integers s, t such that $1 = sa + tb$.*
2. *If $d = \gcd(a, b)$, then $\gcd\left(\frac{a}{d}, \frac{b}{d}\right) = 1$.*

Corollary 4.5 (Euclid's Lemma). *Suppose that a, b are coprime integers and that $a \mid bc$, for some integer c. Then $a \mid c$.*

Example 4.6. *(1) Note that $-2 \mid 16 = (-2) \cdot (-8)$. Hence, by Lemma 4.2, we have that* $\gcd(-2, 16) = 2$.

(2) We have $3 \cdot 17 + (-5) \cdot 10 = 1$. *Hence* 17 *and* 10 *are coprime. Furthermore* $3 \cdot 34 + (-5) \cdot 20 = 2$ *and as* $2 \mid 34$ *and* $2 \mid 20$ *we must have that* $\gcd(34, 20) = 2$.

Theorem 4.7. *(The Euclidean Algorithm) Let $a, b \in \mathbb{Z}$, not both zero and $b \nmid a$. Set $r_{-1} := a$ and $r_0 := b$ and apply the division algorithm successively to r_{k-1} and r_k, for $k \geq 0$ to obtain integers q_k and r_{k+1}, where $0 \leq r_{k+1} < |r_k|$, until $r_{n+1} = 0$, for some $n \geq 0$. That means we have*

$$\begin{array}{lll}
r_{-1} & = q_0 r_0 + r_1, & \text{where } 0 < r_1 < |r_0| \\
r_0 & = q_1 r_1 + r_2, & \text{where } 0 < r_2 < r_1 \\
r_1 & = q_2 r_2 + r_3, & \text{where } 0 < r_3 < r_2 \\
\vdots & \quad\vdots & \\
r_{n-2} & = q_{n-1} r_{n-1} + r_n, & \text{where } 0 < r_n < r_{n-1} \\
r_{n-1} & = q_n r_n + r_{n+1}, &
\end{array}$$

Then $\gcd(a, b) = r_n$.

Proof: As the sequence $(r_k)_{k \geq 1}$ is positive and decreasing, the above process must terminate eventually. In particular r_n exists and $r_n > 0$, by construction. Working through the equations in reverse order one checks that r_n divides both a and b. Finally working through the equations in forward order one checks that every $c \in \mathbb{Z}$ which divides a and b, also divides r_n. In particular $c \leq r_n$, by Lemma 3.3(5). Overall it follows that $\gcd(a, b) = r_n$.

□

Corollary 4.8. (1) $\gcd(a, b) = \gcd(r_k, r_{k+1})$, *for* $k \geq 0$.

(2) *There are $x_i, y_i \in \mathbb{Z}$ such that $r_i = x_i a + y_i b$, for all $i = -1, 0, 1 \ldots, n$, such that $x_{-1} = y_0 = 1$, $y_{-1} = x_0 = 0$ and*

$$x_{i+1} = x_{i-1} - q_i x_i, \qquad \text{and} \qquad y_{i+1} = y_{i-1} - q_i y_i,$$

for all $i = 0, 1, \ldots, n-1$. In particular $\gcd(a, b) = x_n a + y_n b$.

Example 4.9. *Compute* $\gcd(7128, 5148)$.

i	q_i	r_i	x_i	y_i
-1		7128	1	0
0	1	5148	0	1
1	2	1980	1	-1
2	1	1188	-2	3
3	1	792	3	-4
4	2	396	-5	7
5		0		

Thus $\gcd(7128, 5148) = 396$ *and* $396 = (-5) \cdot 7128 + 7 \cdot 5148$. *Furthermore* $\gcd(1980, 1188) = 396$

Remark 4.10. *A* **Diophantine equation** *is an equation in one or more unknowns with integer coefficients, where we are only interested in integer solutions. Let* $a, b, c \in \mathbb{Z}$ *be given. Then* $ax + by = c$ *is a Diophantine equation with two unknown* x *and* y. *For example one can see that* $6x + 4y = 10$ *has a solution* $(x, y) = (1, 1)$, *while* $6x + 4y = 5$ *has no integer solutions.*

Theorem 4.11. *Set* $d = \gcd(a, b)$. *The equation* $ax + by = c$ *is has integer solutions if and only if* $d \mid c$. *If* (x_0, y_0) *is any solution, then the solution set is*

$$\{(x_0 + (b/d)t, y_0 - (a/d)t) : t \in \mathbb{Z}\}.$$

Example 4.12. *Consider the equation* $7128x + 5148y = 792$. *In Example 4.9 we found that* $\gcd(7128, 5148) = 396$. *As* $396 \mid 792 = 2 \cdot 396$, *our equation has a integer solution. Furthermore in the example we found that*

$$396 = (-5) \cdot 7128 + 7 \cdot 5148,$$

and thus $x = -10, y = 14$ *is a solution for our equation. As* $7128/396 = 18$ *and* $5148/396 = 13$, *the set of all solutions is given by*

$$\{(-10 + 13t, 14 - 18t) : t \in \mathbb{Z}\}$$

That means, for $t = -2$ *we get* $(x, y) = (-36, 50)$, *which then is also a solution to our equation.*

In fact, let us find all solutions (x, y) *such that* $15 < x + y < 25$. *As* $x = -10 + 13t$ *and* $y = 14 - 18t$ *we have* $x + y = 4 - 5t$. *Now*

$$15 < x + y < 25 \Leftrightarrow 15 < 4 - 5t < 25 \Leftrightarrow t \in \{-3, -4\}.$$

Hence $(x, y) \in \{(-49, 68), (-62, 86)\}$.

5. The Fundamental Theorem of Arithmetic

Definition 5.1. *An integer $p > 1$ is called* **prime** *if its only positive divisors are 1 and p.*

Example 5.2. *(1) The integer 2 is prime, because assume that $n \mid 2$, where $n \geq 1$. Then $n \leq 2$, by Lemma 3.3(5). Hence $n \in \{1, 2\}$. Similarly we observe that if $n \mid 3$, then $n \in \{1, 2, 3\}$. As $3 = 1 \cdot 2 + 1$, it follows by the Division Algorithm that $2 \nmid 3$. In particular, 3 is prime. Next $4 = 2 \cdot 2$ is not prime. And so on...*

(2) (Sieve of Eratosthenes) Given $n \in \mathbb{N}$ let us find all primes in $S := \{2, \ldots, n\}$. Take $m \in S$. Either m is prime or $m = p \cdot r$, for some $p, r \in S$. We can choose p minimal in S by the Well-Ordering Principle. Then $p \leq r$. Now $p^2 \leq p \cdot r = m$ and so $p \leq \sqrt{m} \leq \sqrt{n}$. Overall m is prime or m is divisible by an integer p with $2 \leq p \leq \sqrt{n}$.

Let $n = 100$. Then $\sqrt{100} = 10$. Eliminating all proper multiples of $\{2, \ldots, 10\}$ in S gives all the primes in S. They are

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97$$

Lemma 5.3. *For integers a, b, let p is a prime divisor of ab. Then $p \mid a$ or $p \mid b$.*

Proof. If $p \mid a$, there is nothing to prove. So assume that $p \nmid a$. Then $\gcd(p, a) = 1$. We deduce from Euclid's Lemma that $p \mid b$. □

Corollary 5.4. *Let $p, a_1, a_2, \ldots, a_k$ be integers, where p is prime. If $p \mid (a_1 a_2 \ldots a_k)$ then $p \mid a_i$, for some $i \in \{1, \ldots, k\}$.*

Corollary 5.5. *Let $p, q_1, q_2, \ldots, q_k$ be prime integers. If $p \mid (q_1 q_2 \ldots q_k)$ then $p = q_i$, for some $i \in \{1, \ldots, k\}$.*

Theorem 5.6 (Fundamental Theorem of Arithmetic). *Given $n \in \mathbb{Z}$, non-zero, there exist $\epsilon \in \{\pm 1\}$ and primes $p_1, \ldots, p_k$ such that*

$$n = \epsilon \cdot p_1 p_2 \ldots p_k.$$

Moreover, this expression is unique up to a permutation of the factors $p_1, \ldots, p_k$.

Proof: W.l.o.g we assume $n \geq 1$. The statement holds for $n = 1$. Next assume the statement holds for $1, \ldots, n - 1$. If n is prime, we are done. Otherwise, there is some positive divisor, say m, of n, such that $m \notin \{1, n\}$. Then $n = m \cdot r$, for some $r \in \mathbb{Z}$. Note that $1 < m, r < n$. By induction m and r are products of prime integers, and thus so is n. The uniqueness of this expression can be shown using Corollary 5.5.

□

Example 5.7. *(1)* $-12 = (-1) \cdot 2^2 \cdot 3$
(2) $21780 = 2 \cdot 10890 = 2 \cdot 5445 = 2^2 \cdot 3 \cdot 1815 = 2^2 \cdot 3^2 \cdot 605 = 2^2 \cdot 3^2 \cdot 5 \cdot 121$
$= 2^2 \cdot 3^2 \cdot 5 \cdot 11^2$
(3) Let's find the prime factorisation of 2369*. Since* $\sqrt{2369} < 49$*, unless* 2369 *is prime, there must be a prime factor of* 2369 *that is* ≤ 48*. By checking* $2, 3, 5, 7, 11, 13, 17, 19, 23$ *we see that* $23 \mid 2369$*, as* $2369 = 23 \cdot 103$*. As* $\sqrt{103} < 11$*, and* $2, 3, 5, 7$ *do not divide* 103 *we conclude that* 103 *is prime. Therefore* $2369 = 23 \cdot 103$ *is the prime factorisation of* 2369*.*

Corollary 5.8. *There are an infinite number of prime integers.*

Proof: Suppose that $p_1, \ldots, p_n$ are all the prime integers. Then $p_1 p_2 \ldots p_n + 1$ is an integer which is not divisible by any of the p_i. Hence there must be further prime integers.

□

Remark 5.9. *Let* $a, b \in \mathbb{Z}$ *be non-zero and let* $p_1, \ldots, p_n$ *be a complete list of prime numbers dividing* a *and/or* b*. Furthermore let*

$$a = p_1^{r_1} \cdots p_n^{r_n}, \quad \textit{and} \quad b = p_1^{s_1} \cdots p_n^{s_n},$$

be the respective prime factorisations of both a *and* b*. (Note that* $r_j, s_j \geq 0$*, for all* $j = 1, \ldots, n$*, but some might be zero.) Then*

$$\gcd(a, b) = p_1^{\min\{r_1, s_1\}} \cdots p_n^{\min\{r_n, s_n\}}$$

For instance we have $21780 = 2^2 \cdot 3^2 \cdot 5 \cdot 11^2$*, and* $15400 = 2^3 \cdot 5^2 \cdot 7 \cdot 11$*. Then*

$$\begin{aligned} \gcd(21780, 15400) &= 2^{\min\{2,3\}} \cdot 3^{\min\{2,0\}} \cdot 5^{\min\{1,2\}} \cdot 7^{\min\{0,1\}} \cdot 11^{\min\{2,1\}} \\ &= 2^2 \cdot 5 \cdot 11 = 220 \end{aligned}$$

6. Modular Arithmetic and Chinese Remainder Theorem

Definition 6.1. *Let a, b, n be integers with $n \geq 1$. We say a* **is congruent to** *b* **modulo** *n and write $a \equiv b \mod n$, if $n \mid (a-b)$.*

Example 6.2. *(1)* $16 \equiv 5 \mod 11$*, since* $11 \mid (16-5)$*.*
(2) $23 \not\equiv 17 \mod 5$*, since* $5 \nmid 23-17=6$*.*
(3) $a \equiv b \mod 1$*, for all* $a, b \in \mathbb{Z}$*, since* $1 \mid (a-b)$*.*
(4) $a \equiv b \mod 2$*, if and only if both a and b are even or both are odd.*

Lemma 6.3. *Let a, n be integers, with $n \geq 1$. Then there exists a unique $r \in \{0, \ldots, n-1\}$ such that $a \equiv r \mod n$. We call r* **residue of** *a* **modulo** *n*

Proof: Observe that $a \equiv r \mod n$ if and only if $n \mid (a-r)$ if and only if there is some $q \in \mathbb{Z}$ such that $a - r = qn$, that is, $a = qn + r$. In particular the existence and uniqueness of $r \in \{0, 1, \ldots, n-1\}$ follows from the Division algorithm.

□

Lemma 6.4. *Let a, b, c, d, n be integers, with $n \geq 1$. Suppose $a \equiv b \mod n$ and $c \equiv d \mod n$. Then*

(1) $a + c \equiv b + d \mod n$
(2) $ac \equiv bd \mod n$
(3) $a^k \equiv b^k \mod n$*, for all integers* $k \geq 0$

Proof: By assumption $n \mid (a-b)$ and $n \mid (c-d)$. Then, by Lemma 3.3 (7), we have $n \mid (a-b+c-d) = ((a+c)-(b+d)$. This gives (1). Also, by Lemma 3.3 (7), we have $n \mid (a-b)c + (c-d)b = ac - db$. This gives (2). Finally part (3) follows from (2).

□

Example 6.5. *(1) What is* $3^{20} \mod 41$*? We have*

$$3^2 = 9 \equiv 9 \mod 41$$
$$3^4 = (3^2)^2 \equiv 9^2 = 81 \equiv -1 \mod 41$$
$$3^8 = (3^4)^2 \equiv (-1)^2 = 1 \equiv 1 \mod 41$$
$$3^{16} = (3^8)^2 \equiv= 1^2 = 1 \equiv 1 \mod 41$$

Now $3^{20} = 3^{16} \cdot 3^4 \equiv 1 \cdot (-1) = -1 \mod 41$*, or in other words* $41 \mid 3^{20} + 1$*.*

Alternatively we have $3^{20} = (3^4)^5 = 81^5 \equiv (-1)^5 = -1 \mod 41$. *Thus again we get* $3^{20} \equiv -1 \mod 41$.

(2) What is the the remainder of $1! + 2! + 3! + 4! + \ldots + 100!$ *upon division by* 12*? Observe that* $12 = 3 \cdot 4$ *divides* $k!$ *for all* $k \geq 4$. *Hence* $k! \equiv 0 \mod 12$, *for all* $k \geq 4$. *Now*

$$1!+2!+3!+4!+\ldots+100! \equiv 1!+2!+3!+0+\ldots+0 = 1!+2!+3! = 9 \mod 12.$$

Remark 6.6. *Let* $a \geq 1$ *be an integer. Furthermore suppose that, read from the left, the digits of* a *are* $d_n, d_{n-1}, \ldots, d_1$, *that is,*

$$a = \sum_{k=1}^{n} d_k \cdot 10^{k-1}.$$

For instance $a = 12375 = 1 \cdot 10000 + 2 \cdot 1000 + 3 \cdot 100 + 7 \cdot 10 + 5 \cdot 10^0$.

(1) Divisibility by 2*: We have* $10^0 = 1 \equiv 1 \mod 2$, *and* $10 \equiv 0 \mod 2$. *By Lemma 6.4 (3) we now have* $10^k \equiv 0^k = 0 \mod 2$, *for all* $k \geq 1$. *Then*

$$a \equiv \sum_{k=1}^{n} d_k \cdot 10^{k-1} \equiv d_1 \mod 2.$$

Hence $a \equiv 0 \mod 2$ *iff* $d_1 \equiv 0 \mod 2$, *or in other words*

$$2 \mid a \text{ iff } 2 \mid d_1.$$

For instance, since $2 \nmid 5$ *we have* $2 \nmid 12375$.

(2) Divisibility by 11*: We have* $10^0 = 1 \equiv 1 \mod 11$, *and* $10 \equiv -1 \mod 11$. *By Lemma 6.4 (3) we now have* $10^k \equiv 1 \mod 11$, *for all even* $k \geq 0$, *and* $10^k \equiv -1 \mod 11$, *for all odd* $k \geq 1$. *Then*

$$a \equiv \sum_{k=1}^{n} d_k \cdot 10^{k-1} \equiv (d_1 + d_3 + \ldots) - (d_2 + d_4 + \ldots) \mod 11$$

Hence $a \equiv 0 \mod 11$ *iff* $(d_1 + d_3 + \ldots) - (d_2 + d_4 + \ldots) \equiv 0 \mod 11$, *or in other words*

$$11 \mid a \text{ iff } 11 \mid (d_1 + d_3 + \ldots) - (d_2 + d_4 + \ldots).$$

For instance, since $11 \mid (5 + 3 + 1) - (7 + 2) = 0$ *we have* $11 \mid 12375$.

Lemma 6.7. *Let* $a, b, c, n \in \mathbb{Z}$ *and* $n \geq 1$ *such that* $ca \equiv cb \mod n$. *Then* $a \equiv b \mod n/d$, *where* $d = \gcd(c, n)$.

Proof: By assumption, there is some $r \in \mathbb{Z}$ such that $nr = (ac - bc) = (a-b)c$. Also $n = d \cdot (n/d)$ and $c = d \cdot (c/d)$. Thus $(n/d) \mid (a-b) \cdot (c/d)$. But $\gcd(n/d, c/d) = 1$, by Corollary 4.4 (2). Hence $(n/d) \mid (a-b)$, by Euclid's lemma. In particular $a \equiv b \mod (n/d)$.

□

Definition 6.8. *Let $n \geq 1$ and a and b be integers. An equation of the form $ax \equiv b \mod n$ is called a* **linear congruence**. *A* **solution** *of such a linear congruence is any integer x_0 such that $ax_0 \equiv b \mod n$. We say two solutions x_1 and x_2 are* **congruent** *if $x_1 \equiv x_2 \mod n$.*

Example 6.9. *Consider the linear congruence $3x \equiv 9 \mod 12$. Clearly $x_1 = 3$ is a solution. Also note that $3 \cdot (-9) = -27 \equiv 9 \mod 12$. Hence $x_2 = -9$ is a solution too. But since $3 \equiv -9 \mod 12$, they are congruent solutions.*

Theorem 6.10. *The linear congruence $ax \equiv b \mod n$ has a solution if and only if $d \mid b$, where $d = \gcd(a,n)$. In this case there are exactly d incongruent solutions modulo n, which are given by*

$$x_0, x_0 + (n/d), x_0 + 2 \cdot (n/d), \ldots, x_0 + (d-1) \cdot (n/d).$$

Proof: omitted

□

Example 6.11. *(1) Consider the linear congruence $9x \equiv 21 \mod 30$. Since $\gcd(9,30) = 3$ and $3 \mid 21$, there are exactly 3 incongruent solutions modulo 30. As the $\gcd(3,30) = 3$, it follows from Lemma 6.7 that $3x \equiv 7 \mod 10$. Since $\gcd(3,10) = 1$ this linear congruence has a unique solution modulo 10. Note that if we multiply $3x \equiv 7 \mod 10$ by 7, we get $21x \equiv 49 \mod 10$, which implies that $x \equiv 9 \mod 10$.*

Now $x = 9$ is a solution of $9x \equiv 21 \mod 30$. Thus its three incongruent solutions are given by $9 + (n/d) \cdot t$, where $t = 1,2,3$. Hence the solutions modulo 30 are $9, 19, 29$.

(2) What is $23^{91} \mod 33$? Let $23^{91} \equiv x \mod 33$. We have

$$\begin{aligned} 23^{91} \equiv x \mod 33 &\Leftrightarrow 33 \mid 23^{91} - x \Leftrightarrow 3 \text{ and } 11 \text{ divide } 23^{91} - x \\ &\Leftrightarrow 23^{91} \equiv x \mod 3 \text{ and } 23^{91} \equiv x \mod 11 \end{aligned}$$

As $23 \equiv -1 \mod 3$ we have $x \equiv 23^{91} \equiv (-1)^{91} = -1 \mod 3$. As $23 \equiv 1 \mod 11$ we have $x \equiv 23^{91} \equiv 1^{91} = 1 \mod 11$. This leads to the system of *linear congruences*

$$x \equiv -1 \mod 3$$
$$x \equiv 1 \mod 11$$

What is x?

Theorem 6.12. *(Chinese Remainder Theorem) Let $n_1, \ldots, n_r$ be positive, pairwise coprime integers, and let $a_1, \ldots, a_r$ be integers. Then the system of linear congruences*

$$\begin{aligned} x &\equiv a_1 \mod n_1 \\ &\vdots \\ x &\equiv a_r \mod n_r \end{aligned}$$

has a simultaneous solution, which is unique modulo $n := n_1 \cdot \ldots \cdot n_r$. This solution is given by

$$\bar{x} = a_1 N_1 x_1 + \ldots + a_r N_r x_r,$$

where $N_k := \dfrac{n}{n_k} = n_1 \cdot n_{k-1} \cdot n_{k+1} \cdot n_r$ and x_k is a solution of $N_k x \equiv 1 \mod n_k$, for all $k = 1, \ldots, r$.

Proof: omitted.

□

Example 6.13. *(1) Let us complete Example 6.11(2). We have the system*

$$x \equiv -1 \mod 3$$
$$x \equiv 1 \mod 11$$

Then $n_1 = 3$, $n_2 = 11$, $n = n_1 \cdot n_2 = 33$, $N_1 = \frac{n}{n_1} = 11$ and $N_2 = \frac{n}{n_2} = 3$.

As $11 \equiv -1 \mod 3$ we have $11x \equiv 1 \mod 3 \Leftrightarrow -x \equiv 1 \mod 3$. Thus $x \equiv -1 \mod 3$, and so $x_1 = -1$ is a solution of $11x \equiv 1 \mod 3$.

Next we look for a solution x_2 of $3x \equiv 1 \mod 11$. Since $4 \cdot 3 = 12 \equiv 1 \mod 11$ we multiply the congruence equation by 4. Then $x \equiv 12x \equiv 4(3x) \equiv 4 \mod 3$. Thus $x_2 = 4$ is a solution of $3x \equiv 1 \mod 11$.

Overall $\bar{x} = a_1 N_1 x_1 + a_2 N_2 x_2 = (-1) \cdot 11 \cdot (-1) + 1 \cdot 3 \cdot 4 = 11 + 12 = 23$ is a simultaneous solution to the given system. Finally this shows that $23^{91} \equiv 23 \mod 33$

(2) Which is the smallest positive number that leaves remainders $2, 3, 2$ *when divided by* $3, 5, 7$, *respectively? That means we look for a solution of the system*

$$x \equiv 2 \mod 3, \quad x \equiv 3 \mod 5, \quad x \equiv 2 \mod 7.$$

Note that $3, 5, 7$ *are pairwise coprime. Hence we can expect a solution. We have* $n = 3 \cdot 5 \cdot 7 = 105$ *and* $N_1 = 5 \cdot 7 = 35$, $N_2 = 3 \cdot 7 = 21$, $N_3 = 3 \cdot 5 = 15$. *This leads to the linear congruences*

$$35x \equiv 1 \mod 3, \quad 21x \equiv 1 \mod 5, \quad 15x \equiv 1 \mod 7,$$

which are equivalent to

$$2x \equiv 1 \mod 3, \quad x \equiv 1 \mod 5, \quad x \equiv 1 \mod 7.$$

So $x_1 = 2$, $x_2 = 1$ *and* $x_3 = 1$ *are their respective solutions. Hence our system has the solution*

$$\begin{aligned} \bar{x} &= a_1 N_1 x_1 + a_2 N_2 x_2 + a_3 N_3 x_3 = 2 \cdot 35 \cdot 2 + 3 \cdot 21 \cdot 1 + 2 \cdot 15 \cdot 1 \\ &= 140 + 63 + 30 = 233 \equiv 23 \mod 105. \end{aligned}$$

So all numbers of the form $23 + 105t$, *for* $t \in \mathbb{Z}$ *solve our system, but* 23 *is the smallest positive such number.*

7. The Real Numbers and the Complex Numbers

Observe that the rational numbers can now be described as

$$\mathbb{Q} = \left\{ \frac{a}{b} : a, b \in \mathbb{Z},\ b \neq 0,\ \gcd(a,b) = 1 \right\}.$$

Lemma 7.1. *There is no $x \in \mathbb{Q}$ such that $x^2 = 2$, that is, the number $\sqrt{2}$ is not a rational number.*

Proof: Let $x \in \mathbb{Q}$ such that $x^2 = 2$. We write $x = a/b$, where $a, b \in \mathbb{Z}$, $b \neq 0$ and a, b are coprime. Then $a^2 = 2b^2$. Hence 2 divides a and thus 4 divides $a^2 = 2b^2$. Now 2 divides b contradicting that a and b are coprime.

□

Remark 7.2. (Real Numbers)
The above Lemma implies that equation $x^2 = 2$ has not solution in $\mathbb{Q}$. Thus one introduces the **real numbers** *as the limit of converging sequences of rational numbers. For instance the sequences $(x_n)_{n\geq 1}$, with $x_1 = 1$ and $x_{n+1} = 4/(x_n + (2/x_n))$, for $n \geq 2$, converges to $\sqrt{2}$. The real numbers are 'ordered' in the sense that $x \leq y$ or $y \leq x$, for any pair x, y of real numbers. The real numbers can be represented on the number line.*

Remark 7.3. (Complex Numbers)
The equation $x^2 = -1$, have no solutions $x \in \mathbb{R}$. This leads to the introduction of the **complex number**, *denoted by $\mathbb{C}$, where*

$$\mathbb{C} := \{a + bi : \ a, b \in \mathbb{R}\},$$

where we define $i^2 = -1$. We say $a + bi \neq c + di$, for $a, b, c, d \in \mathbb{R}$, unless $a = c$ and $b = d$. So, $4 - i \neq 2 + 3i$. For complex numbers $z = a + bi$ and $w = c + di$ we define an addition and multiplication:

$$z + w := (a + c) + (b + d)i, \quad zw := (ac - bd) + (ad + bc)i.$$

For instance

$$(2 + i) + (-1 + 2i) = 1 + 3i \text{ and } (2 + i)(-1 + 2i) = -4 + 3i.$$

These operations are commutative and associative, they have inverses and multiplication distributes over addition.

Complex numbers can be represented in the **complex plane**, *where the x-axis represents the complex numbers with zero imaginary part and the y-axis represents the 'pure' complex numbers with zero real part. Then we can identify z with the point (a, b) in the plane.*

Definition 7.4. *Let $z = a + bi$ be a complex number. Then*

1. $\mathrm{Re}(z) := a$ *is called* **real part** *of z*
2. $\mathrm{Im}(z) := b$ *is called* **imaginary part** *of z*
3. $\overline{z} := a - bi$ *is called* **complex conjugate** *of z*
4. $|z| := \sqrt{a^2 + b^2}$ *is called* **modulus** *or length of z*

Remark 7.5. *(1) There is a copy of $\mathbb{R}$ embedded in $\mathbb{C}$, as we can identify $x \in \mathbb{R}$ with $z := x + 0i \in \mathbb{C}$. The zero element of the complex numbers is $0 + 0i$ and the one element is $1 + 0i$. Unlike $\mathbb{R}$, the complex numbers are not ordered.*

(2) Addition of complex numbers is like vector addition. Furthermore, for $z = a + bi$, the additive inverse is

$$-z = (-a) + (-b)i = -a - bi.$$

(3) Note that $z = a + bi$ is the complex conjugate of $\overline{z}$. Next one checks that $z\overline{z} = (a^2 + b^2) + 0i$. In particular, $z\overline{z} \in \mathbb{R}$ and $z\overline{z} = |z|^2$. Note that $z\overline{z} = 0$ if and only if $a = b = 0$, that is, $z = 0$. Thus whenever $z \neq 0$ we have $z \cdot (\overline{z}/z\overline{z}) = 1$, that is,

$$z^{-1} = \overline{z}/z\overline{z} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} i.$$

For instance

$$\frac{1}{1-i} = \frac{1}{1-i} \cdot \frac{1+i}{1+i} = \frac{1+i}{2} = \frac{1}{2} + \frac{1}{2} i$$

and

$$\frac{2+3i}{1-i} = (2+3i) \cdot \left(\frac{1}{2} + \frac{1}{2} i\right) = \left(1 - \frac{3}{2}\right) + \left(\frac{3}{2} + 1\right) i = \frac{-1}{2} + \frac{5}{2} i$$

(4) Let $z, w \in \mathbb{C}$. Then $\overline{zw} = \overline{z} \cdot \overline{w}$. From this it follows that $|zw|^2 = (zw)\overline{zw} = (z\overline{z})(w\overline{w}) = |z|^2|w|^2$, and so $|zw| = |z||w|$.

(5) Every complex number z is uniquely identified by $|z|$ and its angle θ, in radians, between the positive real axis and z, traced out counter-clockwise. Recall that for all $x \in \mathbb{R}$ we have

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad \sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \quad \cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$

If we extend these definitions to complex numbers, then we get that

$$\exp(xi) = \cos(x) + \sin(x) i.$$

We write e^{xi} *for* $\exp(xi)$*. Note that* $|e^{xi}| = \cos(x)^2 + \sin(x)^2 = 1$*, for any* $x \in \mathbb{R}$*. This gives the* **exponential form** *of* z *as*

$$z = re^{\theta i},$$

where $r = |z|$ *and* θ *is unique modulo* 2π*.*

Take for instance $z = 2 + 3i$*. Then* $|z| = \sqrt{4+9} = \sqrt{13}$*. Hence*

$$z = \sqrt{13} \cdot \left(\frac{2}{\sqrt{13}} + \frac{3}{\sqrt{13}} i \right) = \sqrt{13} \cdot (\cos(\theta) + \sin(\theta)i),$$

for some $\theta \in \mathbb{R}$*. We calculate* $\theta = \arccos\left(\frac{2}{\sqrt{13}}\right) = 0.983$ rad *or* $\theta = 56.3^\circ$*. Since* $\sin(\theta) = \frac{3}{\sqrt{13}}$*, we have* $z = \sqrt{13} \cdot e^{0.983 \cdot i}$*. What about* $w = 2 - 3i$*. Then* $|w| = \sqrt{4+9} = \sqrt{13}$ *and*

$$w = \sqrt{13} \cdot \left(\frac{2}{\sqrt{13}} - \frac{3}{\sqrt{13}} i \right) = \sqrt{13} \cdot (\cos(\eta) + \sin(\eta)i),$$

for some $\eta \in \mathbb{R}$*. Again we calculate* $\eta = \arccos\left(\frac{2}{\sqrt{13}}\right) = 0.983$ rad *or* $\eta = 56.3^\circ$ *and* $\sin(\eta) = \frac{3}{\sqrt{13}} \neq -\frac{3}{\sqrt{13}}$*. Let's switch the sign of* η*, that is, we set* $\eta := -0.983$ rad*. Then* $\cos(\eta) = \frac{2}{\sqrt{13}}$ *and* $\sin(\eta) = -\frac{3}{\sqrt{13}}$*. Hence* $w = \sqrt{13} \cdot e^{-0.983 \cdot i} = \sqrt{13} \cdot e^{(2\pi - 0.983) \cdot i} = \sqrt{13} \cdot e^{5.3 \cdot i}$*. Here* $\eta = 5.3$ rad *or* $\eta = 303.7^\circ$*.*

(6) Now for complex numbers $z = |z|e^{\theta i}$ *and* $w = |w|e^{\eta i}$ *it follows that*

$$zw = |z|e^{\theta i} \cdot |w|e^{\eta i} = |z||w|e^{\theta i} \cdot e^{\eta i} = |zw|e^{(\theta+\eta)i}.$$

Theorem 7.6 (DeMoivre's Theorem)**.** *For all natural numbers* n *and real numbers* θ *we have*

$$(\cos(\theta) + \sin(\theta)i)^n = \cos(n\theta) + \sin(n\theta)i.$$

Proof: The formula can be verified using induction.

□

Definition 7.7. *Let* $n \geq 1$ *be an integer. We say* $z \in \mathbb{C}$ *is a* n*-th root of unity if* $z^n = 1$*.*

Theorem 7.8. *Let* $n \geq 1$ *be a integer. Then there are* n *distinct* n*-th roots of unity, namely the elements of*

$$\left\{ e^{\frac{2\pi k}{n} i} = \cos\left(\frac{2\pi k}{n}\right) + \sin\left(\frac{2\pi k}{n}\right) i : k = 0, 1, \ldots, n-1 \right\}.$$

Example 7.9. *The three third roots of unity are*

$$\begin{aligned} e^{0i} &= \cos(0) + \sin(0)i = 1 \\ e^{\frac{2\pi}{3}i} &= \cos\left(\frac{2\pi}{3}\right) + \sin\left(\frac{2\pi}{3}\right)i = -\frac{1}{2} + \frac{\sqrt{3}}{2}i \\ e^{\frac{4\pi}{3}i} &= \cos\left(\frac{4\pi}{3}\right) + \sin\left(\frac{4\pi}{3}\right)i = -\frac{1}{2} - \frac{\sqrt{3}}{2}i \end{aligned}$$

8. Set Theory

Definition 8.1. *(1) A* set *is a collection of distinct objects. The objects are called* elements. *If A is a set and x is an element in A, then we write $x \in A$. If x does not belong to A, then we write $x \notin A$.*

(2) Let A and B be sets. We say B is a subset *of A, or B is* contained in *A, and write $B \subseteq A$, if for all $x \in B$ we have $x \in A$. If $B \subseteq A$ and there is some $x \in A$ such that $x \notin B$, then we say B is a* proper subset *of A and write $B \subsetneq A$.*

(3) Two sets A and B are called equal *and we write $A = B$, if they have the same elements, or in other words, $A \subseteq B$ and $B \subseteq A$.*

(4) The set that contains no element is called empty set. *We denoted it by $\emptyset$.*

Remark 8.2. *We can describe a set using set notation. For instance we can list the elements in the set as in $A = \{1, 2, 3\}$ or $\mathbb{N} = \{1, 2, 3, \ldots\}$. Or we may describe a set of elements that satisfy a certain property P, e.g. $\{x \in \mathbb{N} : x \text{ satisfies } P\}$.*

Example 8.3. *(1)* $\mathbb{N} = \{1, 2, 3, \dots\}$ *denotes the set of natural numbers. Then* $10 \in \mathbb{N}$ *and* $1/2 \notin \mathbb{N}$.
(2) $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ *denotes the set of integers.*
(3) $\mathbb{Q} = \{\frac{a}{b} : a, b \in \mathbb{Z}, b \neq 0\}$ *denotes the set of rational numbers.*
(4) $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q}$
(5) $\{x \in \mathbb{N} : x < 5\} = \{1, 2, 3, 4\}$
(6) $\{x \in \mathbb{Z} : x^2 + x = 2\} = \{1, -2\}$
(7) $\{3n : \ n \in \mathbb{N}\}$ *describes the set of natural numbers that are multiples of* 3.
(8) Let A be a set. Then $\emptyset \subseteq A$ and $A \subseteq A$.

Definition 8.4. *Let A and B be sets.*
(1) The union *of A and B is the set*

$$A \cup B := \{x : \ x \in A \text{ or } x \in B\}$$

(2) The intersection *of A and B is the set*

$$A \cap B := \{x : \ x \in A \text{ and } x \in B\}$$

(3) The complement *of B relative to A is the set*

$$A \setminus B := \{x : \ x \in A \text{ and } x \notin B\}$$

Example 8.5. *(1)* $\{1,2,3\} \cup \{3,4,5\} = \{1,2,3,4,5\}$
(2) $\{1,2,3\} \cap \{3,4,5\} = \{3\}$
(3) $\{1,2,3\} \setminus \{3,4,5\} = \{1,2\}$
(4) $\mathbb{N} \cup \{-n : n \in \mathbb{N}\} = \mathbb{Z} \setminus \{0\}$
(5) $\{2n : \ n \in \mathbb{N}\} \cap \{3n : \ n \in \mathbb{N}\} = \{6n : \ n \in \mathbb{N}\}$

Definition 8.6. *If the intersection of two sets A and B is empty, then A and B are called* disjoint*.*

Example 8.7. *(1) The alphabet and $\mathbb{Z}$ are disjoint sets.*
(2) The even and the odd integers are disjoint, that is, $\{2n : \ n \in \mathbb{Z}\} \cap \{2n+1 : \ n \in \mathbb{Z}\} = \emptyset$.

Definition 8.8. *Let I be a set, and for each $i \in I$, let A_i be a set. Then*
(1) $\displaystyle\bigcup_{i \in I} A_i := \{x : \ x \in A_i, \text{ for some } i \in I\}$.
(2) $\displaystyle\bigcap_{i \in I} A_i := \{x : \ x \in A_i, \text{ for all } i \in I\}$.

The set I is called index set*.*

Example 8.9. *(1) Let $A_n := \{0, \ldots, n\}$, for all $n \in \mathbb{N}$. We claim that $\displaystyle\bigcup_{n \in \mathbb{N}} A_n = \mathbb{N}$. Clearly $A_n \subseteq \mathbb{N}$, for all $n \in \mathbb{N}$ and so $\displaystyle\bigcup_{n \in \mathbb{N}} A_n \subseteq \mathbb{N}$. On the other hand, if $x \in \mathbb{N}$, then $x \in A_x$, and so $\mathbb{N} \subseteq \displaystyle\bigcup_{n \in \mathbb{N}} A_n$. This proves the claim.*

(2) Let $A_n := \{k \in \mathbb{N} : k \geq n\}$, for all $n \in \mathbb{N}$. We claim that $\displaystyle\bigcap_{n \in \mathbb{N}} A_n = \emptyset$. Assume $\displaystyle\bigcap_{n \in \mathbb{N}} A_n \neq \emptyset$ and let $x \in \displaystyle\bigcap_{n \in \mathbb{N}} A_n$. Then $x \in A_n$, for all $n \in \mathbb{N}$. Since $A_0 = \mathbb{N}$, it follows that $x \in \mathbb{N}$. But then $x \notin A_{x+1}$, which is a contradiction, and therefore $\displaystyle\bigcap_{n \in \mathbb{N}} A_n = \emptyset$.

Definition 8.10. *Let A be a set. Then*

$$\mathcal{P}(A) := \{X : \ X \subseteq A\},$$

the set of all subsets of A, is called power set *of A.*

Example 8.11. *Let $A = \{0,1,2\}$. Then*

$$\mathcal{P}(A) = \{\emptyset, \{0\}, \{1\}, \{2\}, \{0,1\}, \{0,2\}, \{1,2\}, A\}.$$

9. Probability

- Probability: a mathematical model for chance (random) phenomena.
- random process are everywhere:

(1) Weather prediction and climate modelling,
(2) Physics: theory of gases, quantum mechanics
(3) Actuarial science, insurance
(4) Statistics: procedures for analysing data, especially data that has a random character
(5) Genetics and oncology: model for mutations
(6) Transmission of data affected by noise
(7) Games of chance: tossing a coin, rolling a die, card games,

- randomness implies a lack of predictability, we cannot compute the outcome before the event, but we can try to estimate the likelihood of each possible outcome

Definition 9.1. *A* **statistical experiment** *is a process by which we observe something uncertain. An* **outcome** *is a result of a statistical experiment. The set of all outcomes is denoted by Ω. An* **event** *is a subset of Ω. By* **Events of Interest** *we mean any set $\mathscr{E}$ of events, such that*

(E1) $\emptyset \in \mathscr{E}$,
(E2) $\{\omega\} \in \mathscr{E}$, *for all* $\omega \in \Omega$,
(E3) *If* $A \in \mathscr{E}$, *then* $A^c := \Omega \backslash A \in \mathscr{E}$,
(E4) *If* $A_i \in \mathscr{E}$, *for integers* $n \geq 1$, *then* $\bigcup_{i=1}^{\infty} A_i \in \mathscr{E}$

We call the pair $(\Omega, \mathscr{E})$ the **sample space** *of the experiment.*

Remark 9.2. *If Ω has only finitely many outcomes, then $\mathscr{E}$ may contain all subsets of Ω, that is, $\mathscr{E}$. If Ω is infinite, then we have to restrict $\mathscr{E}$ to "certain" subsets.*

Example 9.3. (1) *Experiment: "Rolling a die"*
Outcome: "Throwing a six"
Sample space: $\Omega = \{1, 2, 3, 4, 5, 6\}$
Event: $A =$*"Throwing an even number". Then* $A = \{2, 4, 6\}$

(2) *Experiment: "Tossing a coin",*
Sample space: $\Omega = \{heads, tails\}$
Event: $B =$*"Throwing heads". Then* $B = \{heads\}$

(3) *Experiment: "Picking from a full stack of cards"*
Sample space: $\Omega = \{$*2 of spades, 3 of spades*$, \ldots,$ *ace of spades,*
2 of hearts$, \ldots,$ *ace of clubs*$\}$
Event: $C =$*"Picking hearts". Then* $C = \{$*2 of hearts*$, \ldots,$ *ace of hearts*$\}$

(4) *Experiment: "Picking two balls from an urn with red and blue balls"*
Sample space: $\Omega = \{RR, RB, BR, BB\}$
Event: $D =$*"Picking two different colours". Then* $D = \{RB, BR\}$
Event: $E =$*"Picking blue the second time". Then* $D = \{RB, BB\}$
Event: $F = D \cap E = \{RB\}$

(5) *Experiment: "Cycling through 3 traffic lights, which are red or green"*
Sample space: $\Omega = \{RRR, RRG, RGR, GRR, RGG, GRG, GGR, GGG\}$
Event: $G =$*"Stopping at most once. Then* $G = \{RGG, GRG, GGR, GGG\}$

We want to be able to talk about the relative likelihood of "events of interest" in a sample space.

Definition 9.4. *Let* $(\Omega, \mathscr{E})$ *be the sample space of an experiment. A* **probability measure** *is a function* $P : \mathscr{E} \to \mathbb{R}$*, such that the following* **axioms of probability** *hold:*

(AP1) $P(\Omega) = 1$*;*
(AP2) *If* $A \in \mathscr{E}$ *then* $P(A) \geq 0$*;*
(AP3) *If* A *and* B *are disjoint events in* $\mathscr{E}$*, then* $P(A \cup B) = P(A) + P(B)$*.*

For any outcome $\omega \in \Omega$*, we set* $P(\omega) := P(\{\omega\})$*. We call the triplet* $(\Omega, \mathscr{E}, P)$ *a* **probability space***.*

Lemma 9.5. *Let* $(\Omega, \mathscr{E}, P)$ *be a probability space of an experiment and let* $A, B \in \mathscr{E}$ *be events. Then*

(1) $P(A^c) = 1 - P(A)$
(2) $P(\emptyset) = 0$
(3) *If* $A \subseteq B$ *then* $P(A) \leq P(B)$*.*
(4) $P(A) \in [0, 1]$
(5) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
(6) *If* $A = \{\omega_1, \ldots, \omega_n\}$*, then* $P(A) = P(\omega_1) + \cdots + P(\omega_n)$

Proof. (1) We have that A and A^c are disjoint and $\Omega = A \cup A^c$. Hence by (PA1) and (PA3) we get

$$1 = P(\Omega) = P(A \cup A^c) = P(A) + P(A^c).$$

(2) As $\emptyset = \Omega^c$, it follows from (1) and (PA1) that $P(\emptyset)= P(\Omega^c) = 1-P(\Omega) = 0$.

(3) Note that $B = A\cup(B\backslash A)$, where A and $B\backslash A$ are disjoint. Also $P(B\backslash A) \geq 0$, by (PA2). Hence by (PA3) we have

$$P(A) \leq P(A) + P(B\backslash A) = P(A \cup (B\backslash A)) = P(B).$$

(4) We have $P(A) \geq 0$, by (PA2). As $A \subseteq \Omega$, it follows from (3) and (PA1) that $P(A) \leq P(\Omega) = 1$.

(5) Note that $A \cup B = A \cup (B\backslash A)$, where A and $B\backslash A$ are disjoint. Hence, by (PA3),

$$P(A \cup B) = P(A) + P(B\backslash A).$$

Next observe that $B = (A\cap B)\cup(B\backslash A)$, where $(A\cap B)$ and $B\backslash A$ are disjoint. Hence, by (PA3),

$$P(B) = P(A \cap B) + P(B\backslash A).$$

Putting both equations together we get (5).

(6) The sets $\{\omega_1\}, \ldots, \{\omega_n\}$ are pairwise disjoint and thus a repeated application of (PA3) gives the result. □

Lemma 9.6. *Let $\Omega = \{\omega_1, \ldots, \omega_n\}$. Furthermore assume that $P : \Omega \to [0, 1]$, such that*

$$\sum_{i=1}^{n} P(\omega_i) = 1.$$

Then P becomes a probability measure on the sample space $(\Omega, \mathcal{P}(\Omega))$ by setting

$$P(A) = \sum_{\omega\in A} P(\omega)$$

Example 9.7. *(1) A coin is tossed. Hence we have the sample space $(\Omega, \mathcal{P}(\Omega))$, where $\Omega = \{heads, tails\}$. Now $P(heads) = P(tails) = \frac{1}{2}$, gives rise to a probability space $(\Omega, \mathcal{P}(\Omega), P)$. In this case we call the coin* **fair**. *Likewise $Q(heads) = 0.1$ and $Q(tails) = 0.9$ gives rise to a probability space $(\Omega, \mathcal{P}(\Omega), Q)$.*

(2) In a presidential race there are four candidates $C1$, $C2$, $C3$ and $C4$ and polling suggests that their respective chance of winning is 25%, 15%, 30% and 30%. Consider the probability space $(\Omega, \mathcal{P}(\Omega), P)$, where $\Omega = \{C1, C2, C3, C4\}$ and $P(C1) = 0.25$, $P(C2) = 0.15$ and $P(C3) = P(C4) = 0.3$. If A is the event of $C1$ or $C3$ winning, then $P(A) = P(C1) + P(C3) = 0.25 + 0.3 = 0.55$. Note

that $A^c = \{C2, C4\}$ *and so* $P(A^c) = 1 - P(A) = 0.45 = P(C2) + P(C4)$. *If* B *is the event of* $C3$ *or* $C4$ *winning, then* $P(B) = P(C3) + P(C4) = 0.3 + 0.3 = 0.6$. *Note that* $A \cap B = \{C3\}$ *and so* $P(A \cap B) = P(C3) = 0.3$. *Finally,* $A \cup B = \{C1, C3, C4\}$ *and* $P(A \cup B) = P(A) + P(B) - P(A \cap B) = 0.55 + 0.6 - 0.3 = 0.85 = P(C1) + P(C3) + P(C4)$.

Corollary 9.8. *Let* $(\Omega, \mathcal{P}(\Omega), P)$ *be a finite probability space, where each outcome occurs with the same probability. Then for every event* A

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\textit{number of ways } A \textit{ can occur}}{\textit{total number of outcomes}}$$

Example 9.9. *(1) Toss a fair coin, with outcomes "H" and "T", twice. Then* $\Omega = \{HH, HT, TH, TT\}$. *Note that* $|\Omega| = 4$ *and so each outcome has probability* 0.25. *Next let* A *be the event of heads on first toss and* B *the event of heads on first or second toss. Then* $A = \{HH, HT\}$, $B = \{HT, TH, HH\}$. *Now*

$$P(A) = \frac{|A|}{|\Omega|} = \frac{2}{4} = 0.5 \quad \textit{and} \quad P(B) = \frac{|B|}{|\Omega|} = \frac{3}{4} = 0.75.$$

(2) Two fair dice are rolled in succession. Let A *denote the event of rolling an 8 in total. What is* $P(A)$*? Observe that* $\Omega = \{(i,j) \mid 1 \leq i, j \leq 6\}$, *where each outcome* (i,j) *is equally likely, that is,* $P(i,j) = \frac{1}{|\Omega|} = \frac{1}{36}$. *Then*

$$A = \{(2,6), (3,5), (4,4), (5,3), (6,2)\},$$

and so $P(A) = \frac{|A|}{|\Omega|} = \frac{5}{36}$.

10. Combinatorics

Throughout this section we sample k elements from a set of n elements at random such that

(1) the order in which the k elements are sampled is or is not relevant, (**ordered or unordered sampling**)
(2) each sample is or is not placed back in the set, (**sampling with or without replacement**)

Ordered Sampling with Replacement

Each single draw is the same experiment of sampling one element of n. Hence there are a total of n^k different outcomes.

Example 10.1. *(1) How many different functions $f : \{1, \ldots, 50\} \to \{1, \ldots, 100\}$ are there? For each $k \in \{1, \ldots, 50\}$ we have $f(k) \in \{1, \ldots, 100\}$. Hence we draw 50 samples from a set of 100 elements, where the order matters as the first sample is $f(1)$, the second sample is $f(2)$ and so on, and we allow replacement as $f(1)$ may equal $f(2)$, for instance. Overall, there are 100^{50} different functions.*

(2) How many different subsets does $A = \{x_1, \ldots, x_n\}$ have? Given any subset B of A we ask for each x_i whether $x_i \in B$. Hence B corresponds to a unique ordered list of n 'yes/no' entries. Overall there are 2^n different such lists and so A has 2^n subsets.

Ordered Sampling without Replacement

We consider each individual draw as a separate experiment. Then the first experiment has n possible outcomes, which reduces by one, for each subsequent experiment. Hence, overall there are

$$n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}$$

different outcomes.

Example 10.2. *In a school all pupils are born in the years 2005, 2006 and 2007 (no leap years). If k pupils attend a school event, what is the probability of at least two sharing the same birthday, however, not necessarily the same day of birth?*

We label the kids from 1 *to* k *and let* Ω *be the set of all ordered lists of* k *birthdays. Then* $|\Omega| = 365^k$, *where each outcome is equally likely. Next let* A *be the event that at least two pupils share a birthday. Then* A^c *is the event of no two pupils share a birthday. Note that* $P(A) = 1 - P(A^c) = 1 - \frac{|A^c|}{|\Omega|}$.

Next observe that A^c *is precisely those ordered lists in* Ω *that have no repeated birthdays. Hence* $|A^c| = 365 \cdot \ldots \cdot (366 - k) = \frac{365!}{(365-k)!}$. *Thus*

$$P(A) = 1 - \frac{|A^c|}{|\Omega|} = 1 - \frac{365!}{365^k \cdot (365-k)!}$$

How many pupils need to attend the school event for there to be at least a 50 percent chance of a shared birthday? One can calculate that $P(A) \approx 0.4757$, *for* $k = 22$ *and* $P(A) \approx 0.5073$, *for* $k = 23$.

Fun Fact: *At the 2014 Football World Cup there were 32 teams of 23 players each and in exactly 16 teams at least two players with a shared birthday.*

Unordered Sampling without Replacement

Theorem 10.3. *The number of total outcomes in this scenario is*

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

Proof. Let x be the number of total outcomes. For each unordered sampling of k elements without replacement there are $k!$ different ways to order the sampling. Hence $x \cdot k!$ equals the number of ordered samplings of k elements without replacement, which we know equals $\frac{n!}{(n-k)!}$. This gives the result. □

Example 10.4. *(1) We draw three cards at random from a stack of 52 cards. Let* Ω *be the set of all possible combinations ignoring their order. Then*

$$|\Omega| = \binom{52}{3} = \frac{52 \cdot 51 \cdot 50}{6} = 22100.$$

Note that all outcomes are equally likely. Next let A *be the event that exactly one ace has been drawn. This is like drawing one card from the four aces and two cards from the remaining* 48. *Hence*

$$|A| = \binom{48}{2} \cdot \binom{4}{1} = \frac{48 \cdot 47}{2} \cdot 4 = 4512.$$

Now $P(A) = \frac{4512}{22100} = 0.2042$.

(2) A radio DJ wants to play 12 songs over the next hour, 4 each from the 60s, 70s and 80s. How many ways are there to arrange the song list by decade? In other words, the DJ wants to choose 4 of 12 slots to play a 60s song, 4 of the remaining 8 slots to play a 70s song and then fill the final 4 slots with 80s songs. Hence there are

$$\binom{12}{4} \cdot \binom{8}{4} \cdot \binom{4}{4} = 495 \cdot 70 \cdot 1 = 34650$$

different arrangements by decade.

Lemma 10.5. *Let $a, b \in \mathbb{R}$ and $n \geq 1$ an integer. Then*

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Proof. Clearly, every summand of $(a+b)^n$ is a multiple of $a^k b^{n-k}$, for some $k \in \{0, 1, \ldots, n\}$. This multiple is the number of ways that one can pick k times the scalar a from the n factors of $(a+b)^n$. □

Lemma 10.6. Vandermonde's identity*:*

$$\binom{m+n}{k} = \sum_{i=0}^{k} \binom{m}{i} \cdot \binom{n}{k-i}$$

Proof. Let $S = \{a_1, \ldots, a_m, b_1, \ldots, b_n\}$. Note that the number of subsets of S with k elements is $\binom{m+n}{k}$. On the other hand any such subset will contain i elements in $\{a_1, \ldots, a_m\}$ and $k - i$ elements in $\{b_1, \ldots, b_n\}$, for some $i \in \{0, 1, \ldots, k\}$. But there are $\binom{m}{i} \cdot \binom{n}{k-i}$ for each i. This proves the claim. □

Unordered Sampling with Replacement

Theorem 10.7. *The number of total outcomes in this scenario is*

$$\binom{n+k-1}{k} = \binom{n+k-1}{n-1}.$$

Proof. Assume we sample from the set $\{x_1, \ldots, x_n\}$. For each $i \in \{1, \ldots, n\}$, let k_i denote the number of times we have drawn x_i. Then $k_1 + \ldots + k_n = k$ and each such n-tuple $(k_1, \ldots, k_n)$ represents a unique outcome of the sampling experiment. Next we replace k_i by k_i vertical bars. Now the sum $k_1 + \ldots + k_n$ corresponds to a unique sequence of k bars and $n - 1$ plus signs. For instance, $3 + 0 + 1$ corresponds to $||| + +|$. Such a sequence has $n + k - 1$ spots, from which we sample (unordered and without replacement) k spots to place the bars. □

Example 10.8. *Twenty people take the bus from Galway to Dublin. The bus stops five times and the bus driver records the number of people getting off at each stop. How many different possibilities exist? Let* $S1$ *to* $S5$ *denote the five stops. For each passenger we draw a sample from the set* $\{S1, \ldots, S5\}$*, where we allow replacement but the order does not matter, (that is, the order in which each group gets off the bus at a particular stop). Hence there are*

$$\binom{5+20-1}{20} = \binom{24}{4} = 10,626$$

possible outcomes.

11. Conditional Probability

Example 11.1. *Consider an urn with five red and five blue balls. We pick all ten balls, one at a time, without replacement. Prior to the experiment the chance to pick blue as the i-th ball is $\frac{1}{2}$. However this chance changes as the experiment progresses. Assume we pick a blue first ball. Then the chance to pick blue as the second ball drops to $\frac{4}{9}$. In fact, if A is the event to pick a blue second ball and B is the event to pick a blue first ball, then the chance that A happens, given that B has happened, is identical to the $\frac{P(A\cap B)}{P(B)}$.*

Definition 11.2. *In a finite probability space $(\Omega, \mathcal{P}(\Omega), P)$ let $A, B \subseteq \Omega$ such that $P(B) \neq 0$. Then the* **conditional probability** *of A given B is defined as*

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

The idea is that $P(A \mid B)$ gives the probability that event A occurs, given that we know that B has occurred.

Example 11.3. *We roll a fair die twice, with respective results X_1 and X_2. Assuming that $X_1 + X_2 = 8$, what is the probability of there having been thrown (i) a one, (ii) a two and (iii) a four? Let $\Omega = \{(X_1, X_2) : X_1, X_2 \in \{1, \ldots, 6\}\}$. Also let $B = \{(X_1, X_2) \in \Omega : X_1 + X_2 = 8\}$. Then*

$$B = \{(2,6), (3,5), (4,4), (5,3), (6,2)\},$$

and $P(B) = \frac{5}{36}$. Next, for $i = 1, \ldots, 6$, et A_i be the event that the number i is thrown at least once. Then

$$A_1 \cap B = \emptyset, \quad A_2 \cap B = \{(2,6), (6,2)\} \quad A_4 \cap B = \{(4,4)\}.$$

Hence

$$\begin{aligned} P(A_1 \mid B) &= \frac{P(A_1 \cap B)}{P(B)} = \frac{0}{\frac{5}{36}} = 0 \\ P(A_2 \mid B) &= \frac{P(A_2 \cap B)}{P(B)} = \frac{\frac{2}{36}}{\frac{5}{36}} = \frac{2}{5} \\ P(A_4 \mid B) &= \frac{P(A_4 \cap B)}{P(B)} = \frac{\frac{1}{36}}{\frac{5}{36}} = \frac{1}{5} \end{aligned}$$

Theorem 11.4 (Law of Total Probability)**.** *In a finite probability space $(\Omega, \mathcal{P}(\Omega), P)$ let $B_1, \ldots, B_n$ be disjoint events with $P(B_i) > 0$, for all $i = 1, \ldots, n$ and*

$\Omega = B_1 \cup \cdots \cup B_n$. *Then for any event* A *we have*

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

Proof. Note that $A = (A \cap B_1) \cup \cdots \cup (A \cap B_n)$ is a union or pairwise disjoint sets, and so

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(A \mid B_i)P(B_i).$$

□

Example 11.5. *(1) In a game a player* X *throws a fair die and receives a bag with 100 marbles where* r *are red and the rest blue. If* X *throws* 1, 2 *or* 3, *then* $r = 45$, *if* X *throws* 4 *or* 5, *then* $r = 60$. *If* X *throws* 6, *then* $r = 75$. *Now* X *picks a marble at random from the bag. The player wins if the marble is red. What is the chance of winning?*

Let R *be the event that the chosen marble is red and let* B_r *be the event that the bag contains* r *red marbles. Then* $P(R \mid B_r) = \frac{r}{100}$. *Overall*

$$P(R) = \sum_{r \in \{45,60,75\}} P(R \mid B_r)P(B_r) = \frac{45}{100} \cdot \frac{1}{2} + \frac{60}{100} \cdot \frac{1}{3} + \frac{75}{100} \cdot \frac{1}{6} = 0.55$$

(2) We consider families with two children (girl and/or boy) and assume all outcomes $\{GG, GB, BG, BB\}$ *are equally likely. What is the probability that both children are girls given that (a) the first born is a girl, (b) the family has a girl? Let* $A = \{GG\}$, $B = \{GG, GB\}$ *and* $C = \{GG, GB, BG\}$. *Then*

$$P(A) = \frac{1}{4}, \qquad P(B) = \frac{1}{2}, \qquad P(C) = \frac{3}{4}$$

(a) We want $P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(B)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$.

(b) We want $P(A \mid C) = \frac{P(A \cap C)}{P(C)} = \frac{P(A)}{P(C)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$.

Theorem 11.6 (Bayes' Rule)**.** *In a finite probability space* $(\Omega, \mathcal{P}(\Omega), P)$ *let* A *and* B *be events such that* $P(A) \neq 0$. *Then*

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

Example 11.7. *(1) We continue with Example 11.5(1). What is the probability that bag B_r was chosen, given that the player won, that is, picked a red marble? We ask for $P(B_r \mid R)$. We have*

$$P(B_r \mid R) = \frac{P(R \mid B_r) \cdot P(B_r)}{P(R)} = \frac{\frac{r}{100}}{0.55} \cdot P(B_r) = \frac{r}{55} \cdot P(B_r)$$

Hence

$$P(B_{45} \mid R) = 0.4091, \qquad P(B_{60} \mid R) = 0.3636, \qquad P(B_{75} \mid R) = 0.2273$$

(2) We continue with Example 11.5(2). Let us assume that girls are called Lara with a very small probability $\alpha << 0.1$. Assume that a family has a girl named Lara. What is the probability of the family having two girls. Let L be the event of there being a child named Lara. We want $P(GG \mid L)$. Note that

$$\begin{aligned} &P(L \mid BB) = 0 \\ &P(L \mid GB) = P(L \mid BG) = \alpha \\ &P(L \mid GG) = \alpha + (1-\alpha)\alpha = 2\alpha - \alpha^2 \end{aligned}$$

By the Law of Total Probability we have

$$\begin{aligned} P(L) &= (P(L \mid BB) + P(L \mid GB) + P(L \mid BG) + P(L \mid GG)) \cdot \frac{1}{4} \\ &= (4\alpha - \alpha^2) \cdot \frac{1}{4}. \end{aligned}$$

Now by Bayes' Rule

$$P(GG \mid L) = \frac{P(L \mid GG) \cdot P(GG)}{P(L)} = \frac{(2\alpha - \alpha^2) \cdot \frac{1}{4}}{(4\alpha - \alpha^2) \cdot \frac{1}{4}} = \frac{2-\alpha}{4-\alpha} \approx \frac{1}{2}$$

12. Independence

Definition 12.1. *Let $(\Omega, \mathscr{E}, P)$ be a probability space. We call two events A, B* **independent** *if*

$$P(A \cap B) = P(A) \cdot P(B).$$

Remark 12.2. *Note that A, B are independent if and only if $P(A) = P(A \mid B)$ and $P(B) = P(B \mid A)$.*

Example 12.3. *(1) Suppose we pick at random a number n from the set $\{1, \ldots, 10\}$, that is, each outcome is equally likely. Let A be the event that n is less than 7 and let B be the event that n is even. Then*

$$A = \{1, 2, 3, 4, 5, 6\}, \quad B = \{2, 4, 6, 8, 10\} \text{ and } A \cap B = \{2, 4, 6\},$$

and so

$$P(A) = \frac{6}{10} = 0.6, \quad P(B) = \frac{1}{2} = 0.5 \text{ and } P(A \cap B) = \frac{3}{10} = 0.3.$$

Hence $P(A \cap B) = P(A) \cdot P(B)$ and so A and B are independent.

However if we let A be the event that n is less than 8, then $P(A) = 0.7$ and A and B are dependent, because $P(A \mid B) = \frac{P(A \cap B)}{P(B)} = 0.6$.

(2) A fair coin is tossed twice. The outcome of the two tosses are independent. However, next let two equally good snooker players X and Y face each other in a match. We assume the first frame is a $50/50$ affair. But since losing a frame increases the pressure, the loser's chances of winning the next frame drop by one percent. Next let A be the event that player X wins the first game and B the event that player X wins the second game. Then

$$P(A) = \frac{1}{2}, \quad \text{and} \quad P(B) = \frac{1}{2} \cdot \frac{49}{100} + \frac{1}{2} \cdot \frac{51}{100} = \frac{1}{2},$$

and so $P(A) \cdot P(B) = \frac{1}{4}$. However

$$P(A \cap B) = \frac{1}{2} \cdot \frac{51}{100} = \frac{51}{200} = \frac{1}{4} + \frac{1}{200}.$$

Hence the outcome of successive games is dependent. Also note that

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{1}{2} + \frac{1}{100} = \frac{P(A \cap B)}{P(A)} = P(B \mid A).$$

13. Discrete Random Variables

Definition 13.1. *Let $(\Omega, \mathscr{E}, P)$ be a probability space. A function $X : \Omega \to \mathbb{R}$ is called* **random variable***. We say X is* **discrete** *if its range R_X is a countable set, that is, it is either finite or bijective to the natural numbers.*

Example 13.2. *(1) We toss a fair coin 3 times. Then*

$$\Omega = \{HHH, HHT, HTH, THH, THT, HTT, TTH, TTT\}.$$

Let X be the number of heads per outcome. Then for instance $X(HHH) = 3$ and $X(THT) = 1$. Furthermore $R_X = \{0, 1, 2, 3\}$. Other possible random variables are for instance

1. *Y is the number of tails per outcome; here $R_Y = \{0, 1, 2, 3\}$*
2. *$Z = X - Y$ is the number of heads minus the number of tails; here $R_Z = [-3, 3] \cap \mathbb{Z}$.*

(2) We toss a fair coin until the first heads appears. If X is the number of throws, then $R_X = \{1, 2, 3, \ldots\}$.

(3) Let X be the random variable that gives the exact moment from now in minutes until the next rainfall occurs in a certain place. Then $R_X = \mathbb{N}$.

Definition 13.3. *Let $(\Omega, \mathscr{E}, P)$ be a probability space and $X : \Omega \to \mathbb{R}$ a discrete random variable with range R_X. Then the function $P_X : R_X \to [0, 1]$, where*

$$P_X(x) = P(X = x) = P(\{\omega \in \Omega : X(\omega) = x\}),$$

for every $x \in R_X$, is called the **probability mass function** *(PMF) of X.*

The **cumulative distribution function** *(CDF) of X is defined as*

$$F_X(x) = P(X \le x), \text{ for all } x \in \mathbb{R}.$$

Example 13.4. *(1) We toss a fair coin 3 times. Let X be the number of heads per outcome. Then $R_X = \{0, 1, 2, 3\}$ and the PMF of X is*

$$\begin{aligned}
P_X(0) &= P(X = 0) = P(\{TTT\}) = \frac{1}{8} \\
P_X(1) &= P(X = 1) = P(\{TTH, THT, HTT\}) = \frac{3}{8} \\
P_X(2) &= P(X = 2) = P(\{HHT, HTH, THH\}) = \frac{3}{8} \\
P_X(3) &= P(X = 3) = P(\{HHH\}) = \frac{1}{8}
\end{aligned}$$

Furthermore

$$F_X(x) = \begin{cases} 0, & x \in (-\infty, 0) \\ \frac{1}{8}, & x \in [0, 1) \\ \frac{1}{2}, & x \in [1, 2) \\ \frac{7}{8}, & x \in [2, 3) \\ 1, & x \in [3, \infty) \end{cases}$$

(2) We toss a coin, where $P(H) = p$, for $0 < p < 1$, until the first heads appears. If X is the number of throws, then $R_X = \{1, 2, 3, \ldots\}$. Now

$$P_X(1) = p, \qquad P_X(2) = (1-p)p, \qquad P_X(3) = (1-p)^2 p, \qquad \ldots$$

In fact, we have $P_X(k) = p(1-p)^{k-1}$, for all $k \in R_X$. Hence, if the coin is fair, then $P_X(k) = \frac{1}{2^k}$, for all $k \in R_X$.

Next observe that, for all integers $n \geq 1$,

$$\begin{aligned} F_X(n) &= \sum_{k=1}^{n} P_X(k) = p \cdot \sum_{k=1}^{n} (1-p)^{k-1} = (1 - (1-p)) \cdot \sum_{k=1}^{n} (1-p)^{k-1} \\ &= \sum_{k=1}^{n} (1-p)^{k-1} - \sum_{k=1}^{n} (1-p)^{k} = 1 - (1-p)^n \end{aligned}$$

Lemma 13.5. *Let $(\Omega, \mathscr{E}, P)$ be a probability space with discrete random variable $X : \Omega \to \mathbb{R}$. Then*

(1) $\displaystyle\sum_{x \in R_X} P_X(x) = 1$,

(2) $P(a < X \leq b) = F_X(b) - F_X(a)$, *for real $a < b$.*

Proof. Part (1) follows immediately from Lemma 9.6. For part (2) observe that $P(X \leq b) = P(X \leq a) + P(a < X \leq b)$. □

Example 13.6. *Consider the PMF $P_X(k) = p(1-p)^{k-1}$, for all integers $k \geq 1$, where $0 < p < 1$, from Example 13.4 (2). Then*

$$\sum_{k=1}^{\infty} P_X(k) = \lim_{n \to \infty} \sum_{k=1}^{n} P_X(k) = \lim_{n \to \infty} F_X(n) = \lim_{n \to \infty} (1 - (1-p)^n) = 1$$

Furthermore

$$\begin{aligned} P(10 < X \leq 100) &= F_X(100) - F_X(10) = (1 - (1-p)^{100}) - (1 - (1-p)^{10}) \\ &= (1-p)^{10} - (1-p)^{100} = (1-p)^{10} \cdot \big(1 - (1-p)^{90}\big) \end{aligned}$$

Definition 13.7. *We call discrete random variables* $X_1, \ldots, X_n$ **independent** *if, for all sets* $A_1, \ldots, A_n$ *in* $\mathbb{R}$,

$$P(X_1 \in A_1, \ldots, X_n \in A_n) = P(X_1 \in A_1) \cdot \ldots \cdot P(X_n \in A_n).$$

Example 13.8. *We throw a blue die and red die ten times each independent of each other. Let* X *count the number of red* 6*'s and let* Y *count the number of blue odd numbers. We have*

$$\begin{aligned} P(X < 2, Y > 1) &= P(X < 2) \cdot P(Y > 0) = (P_X(0) + P_X(1)) \cdot (1 - P_Y(0)) \\ &= \left(\left(\frac{5}{6}\right)^{10} + 10 \cdot \frac{1}{6} \cdot \left(\frac{5}{6}\right)^{9}\right) \cdot \left(1 - \left(\frac{1}{2}\right)^{10}\right) \\ &= \left(\frac{5}{6}\right)^{9} \cdot \left(\frac{5}{2}\right) \cdot \left(1 - \left(\frac{1}{2}\right)^{10}\right) = 0.484 \end{aligned}$$

14. Special Random Variables and Distributions

Bernoulli Random Variables. A Bernoulli trial, is an experiment with two possible outcomes, called success and failure, that is, $\Omega = \{success, failure\}$. Assume that success happens with probability p, for some $0 < p < 1$. Next let $X : \Omega \to \mathbb{R}$ such that $X(failure) = 0$ and $X(success) = 1$. Then $R_X = \{0, 1\}$ and

$$P_X(x) = \begin{cases} p, & x = 1 \\ 1-p, & x = 0 \\ 0, & \text{otherwise} \end{cases}$$

We say X is a **Bernoulli Random Variable** / has a **Bernoulli distribution**, with parameter p, and write $X \sim \text{Bernoulli}(p)$.

Example 14.1. *(1) Toss a coin and let X record 1 for heads and 0 for tails. If the coin is fair, then X has a Bernoulli distribution with parameter $\frac{1}{2}$, that is, $X \sim \text{Bernoulli}(\frac{1}{2})$.*

(2) We roll a die and let X record 1 if the number is 5 or 6, and 0 otherwise. Then X is a Bernoulli Random Variable with parameter $p = 1/3$.

Binomial Distribution. Consider an experiment of n independent Bernoulli trials, each with the same probability of success p. Let X be the random variable that records the number of successes. Then $R_X = \{0, 1, \ldots, n\}$. Let $k \in R_X$. Note that there are $\binom{n}{k}$ different ways to record precisely k successes. Furthermore each such outcome occurs with probability $p^k \cdot (1-p)^{n-k}$. Hence

$$P_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k}, & k = 0, 1, \ldots, n \\ 0, & \text{otherwise} \end{cases}$$

We say X has a **Binomial distribution**, with parameters n and p, and write $X \sim \text{Binomial}(n, p)$.

Example 14.2. *(1) Toss a coin 100 times, where head represents a success and let X record the number of successes. If the coin is fair, then X has a Binomial distribution with parameters 100 and $\frac{1}{2}$, that is, $X \sim \text{Binomial}(100, \frac{1}{2})$*

(2) Rolling a die we call the outcomes 5 or 6 a success. We roll four times and let X record the number of successes. Then $X \sim \text{Binomial}(4, \frac{1}{3})$ and

k	0	1	2	3	4
$P_X(k)$	0.1975	0.3951	0.2963	0.0988	0.0123

Geometric Distribution. Consider an experiment where we run independent Bernoulli trials, each with probability of success p, until the first success is observed. Let X count the number of trials, that is, the experiment terminates after X trails. As seen in Example 13.4 (2) we get

$$P_X(k) = \begin{cases} p(1-p)^{k-1}, & k = 1, 2, 3\dots \\ 0, & \text{otherwise} \end{cases}$$

We say X has a **Geometric distribution**, with parameter p, and write $X \sim \text{Geometric}(p)$. For instance, let $p = \frac{1}{3}$,

k	1	2	3	4
$P_X(k)$	$\frac{1}{3} = 0.333$	$\frac{2}{9} = 0.222$	$\frac{4}{27} = 0.148$	$\frac{8}{81} = 0.0987$

Pascal Distribution. In a variant, Bernoulli trials are repeated until there are r successes. Again let X record the number of trials. Then

$$P_X(k) = \begin{cases} \binom{k-1}{r-1} p^r (1-p)^{k-r}, & k = r, r+1, r+2, r+3\dots \\ 0, & \text{otherwise} \end{cases}$$

In this case X has **Pascal distribution**, with parameters r and p, and write $X \sim \text{Pascal}(r, p)$.

Example 14.3. *We throw a fair coin until we have thrown three Heads. If X counts the number of throws needed, then $X \sim \text{Pascal}(3, \frac{1}{2})$. We have*

k	0	1	2	3	4	5	6	7	8
$P_X(k)$	0	0	0	0.125	0.1875	0.1875	0.1563	0.1172	0.082

Poisson Distribution. Let $\lambda > 0$ be real and $X_n \sim \text{Binomial}(n, p_n)$, where $p_n = \frac{\lambda}{n}$, for integers $n \geq 1$. Set $q_n = 1 - p_n$. Then for fixed $k \in \mathbb{N}$ we have

$$P(X_n = k) = \binom{n}{k} p_n^k q_n^{n-k} = \frac{\lambda^k}{k!} \cdot \frac{n(n-1)\dots(n-k+1)}{n^k} \cdot \frac{(1-\lambda/n)^n}{(1-\lambda/n)^k}$$

Now let $n \to \infty$. Then, one can show that,

$$\frac{n(n-1)\dots(n-k+1)}{n^k} \to 1, \quad (1-\lambda/n)^n \to e^{-\lambda}, \quad (1-\lambda/n)^k \to 1.$$

So

$$P(X_n = k) \to \frac{\lambda^k e^{-\lambda}}{k!}.$$

We say a random variable X has **Poisson distribution** with parameter $\lambda > 0$, and write $X \sim \text{Poisson}(\lambda)$, if

$$P_X(k) = \begin{cases} \frac{\lambda^k}{k!} e^{-\lambda}, & k = 0, 1, 2, 3, \dots \\ 0, & \text{otherwise} \end{cases}$$

Lemma 14.4. *The above function is a probability mass function on the set* $\mathbb{N}$.

Proof. We have $\sum_{k=0}^{\infty} P_X(k) = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1$ □

Remark 14.5. *The Poisson distribution can be used to approximate a Binomial distribution for large n and small p. Furthermore the Poisson distribution is used when we count the occurrences of certain events in time or space, knowing that on average they occur* λ *times.*

Example 14.6. *(1) Two dice are rolled* 100 *times, and the number of double sixes is counted by* X*. Each dice roll is a Bernoulli trial with* $p = 1/36$*. So* $X \sim \text{Binomial}(n = 100, p = 1/36)$*. Since n is large relative to p, we can approximate X as a Poisson random variable, with* $\lambda = p \cdot n = \frac{25}{9}$*. Thus*

$$P_X(k) \approx \frac{(\frac{25}{9})^k}{k!} e^{-\frac{25}{9}}$$

We have

k	0	1	2	3	4	5	6	7
Binomial$(X \leq k)$	0.06	0.231	0.472	0.698	0.854	0.94	0.979	0.993
Poisson$(X \leq k)$	0.062	0.235	0.475	0.697	0.851	0.937	0.977	0.992

(2) A small store has ten customers per hour on average. Let X be the number of customers arriving between 10am to 11.30am? What is the probability of there being more than 12, but no more than 16 customers during that time? We assume that X is Poisson distributed with parameter $\lambda = 1.5 \cdot 10 = 15$ *and want* $P(12 < X \leq 16)$*. We have*

$$\begin{aligned} P(12 < X \leq 16) &= \sum_{k=13}^{16} P_X(k) = \sum_{k=13}^{16} \frac{15^k}{k!} e^{-15} \\ &= e^{-15} \cdot \left(\frac{15^{13}}{13!} + \frac{15^{14}}{14!} + \frac{15^{15}}{15!} + \frac{15^{16}}{16!} \right) = 0.3965 \end{aligned}$$

15. Expected Value, Variance, Standard Deviation

Definition 15.1. *Let X be a discrete random variable with range R_X. The* **expected value, mean** *or* **average** *of X, denoted by $\mathbb{E}(X)$, is defined as*

$$\mathbb{E}(X) = \sum_{k \in R_X} k \cdot P_X(k)$$

Example 15.2. *(1) We throw a die, that is, $\Omega = \{1, 2, 3, 4, 5, 6\}$ and let X denote the outcome, that is, $X(\omega) = \omega$. Then*

$$\mathbb{E}(X) = \sum_{k=1}^{6} k \cdot \underbrace{P_X(k)}_{=\frac{1}{6}} = \frac{1}{6} \cdot (1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = 3.5$$

(2) Let $X \sim \text{Bernoulli}(p)$. Then $R_X = \{0, 1\}$ and

$$\mathbb{E}(X) = 0 \cdot P_X(0) + 1 \cdot P_X(1) = 0 \cdot (1 - p) + 1 \cdot p = p$$

(3) Let $X \sim \text{Geometric}(p)$. Then $R_X = \{1, 2, 3, \ldots\}$ and $P_X(k) = p(1-p)^{k-1}$, for $k \in R_X$. Then

$$\mathbb{E}(X) = \sum_{k=1}^{\infty} k \cdot P_X(k) = p \cdot \underbrace{\sum_{k=1}^{\infty} k \cdot (1 - p)^{k-1}}_{=\frac{1}{(1-(1-p))^2}} = \frac{1}{p}$$

(4) Let $X \sim \text{Poisson}(\lambda)$. Then $R_X = \{0, 1, 2, \ldots\}$ and $P_X(k) = \frac{e^{-\lambda}\lambda^k}{k!}$, for $k \in R_X$. Then

$$\begin{aligned}\mathbb{E}(X) &= \sum_{k=1}^{\infty} k \cdot P_X(k) = e^{-\lambda} \cdot \sum_{k=1}^{\infty} \frac{\lambda^k}{(k-1)!} = \lambda \cdot e^{-\lambda} \cdot \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} \\ &= \lambda \cdot e^{-\lambda} \cdot e^{\lambda} = \lambda\end{aligned}$$

Lemma 15.3. *Let X and $X_1, \ldots, X_n$ be random variables and $a, b \in \mathbb{R}$. Then*

(1) $\mathbb{E}(a \cdot X + b) = a \cdot \mathbb{E}(X) + b$
(2) $\mathbb{E}(X_1 + \ldots + X_n) = \mathbb{E}(X_1) + \ldots + \mathbb{E}(X_n)$

Proof. (1) Let $Y = a \cdot X + b$. Then $R_Y = \{a \cdot x + b : \ x \in R_X\}$. Also $P(Y = a \cdot x + b) = P(X = x)$, for all $x \in R_X$. Now

$$\mathbb{E}(a \cdot X + b) = \sum_{y \in \mathbb{R}_Y} y \cdot P_X(y) = \sum_{x \in \mathbb{R}_X} (a \cdot x + b) \cdot P_X(x) = a \cdot \mathbb{E}(X) + b$$

(2) We show the case $n = 2$. First note that

$$\begin{aligned}\mathbb{E}(X_1) &= \sum_{k_1 \in R_{X_1}} k_1 \cdot P(X_1 = k_1) = \sum_{k_1 \in R_{X_1}} k_1 \cdot \sum_{k_2 \in R_{X_2}} P(X_1 = k_1, X_2 = k_2) \\ &= \sum_{\substack{k_1 \in R_{X_1} \\ k_2 \in R_{X_2}}} k_1 \cdot P(X_1 = k_1, X_2 = k_2)\end{aligned}$$

and likewise

$$\mathbb{E}(X_2) = \sum_{\substack{k_1 \in R_{X_1} \\ k_2 \in R_{X_2}}} k_2 \cdot P(X_1 = k_1, X_2 = k_2)$$

Then

$$\begin{aligned}\mathbb{E}(X_1 + X_2) &= \sum_{\substack{k_1 \in R_{X_1} \\ k_2 \in R_{X_2}}} (k_1 + k_2) \cdot P(X_1 = k_1, X_2 = k_2) \\ &= \sum_{\substack{k_1 \in R_{X_1} \\ k_2 \in R_{X_2}}} k_1 \cdot P(X_1 = k_1, X_2 = k_2) + \sum_{\substack{k_1 \in R_{X_1} \\ k_2 \in R_{X_2}}} k_2 \cdot P(X_1 = k_1, X_2 = k_2) \\ &= \mathbb{E}(X_1) + \mathbb{E}(X_2)\end{aligned}$$

□

Example 15.4. *(1) Let $X \sim \text{Binomial}(n, p)$. Then $X = \sum_{i=1}^{n} Y_i$, where $Y_i \sim \text{Bernoulli}(p)$, for all $i = 1, \ldots, n$. Now*

$$\mathbb{E}(X) = \mathbb{E}(Y_1 + \ldots + Y_n) = n \cdot \mathbb{E}(Y_1) = n \cdot p$$

(2) Let $X \sim \text{Pascal}(r, p)$. Then $X = \sum_{i=1}^{r} Y_i$, where $Y_i \sim \text{Geometric}(p)$, for all $i = 1, \ldots, r$. Now

$$\mathbb{E}(X) = \mathbb{E}(Y_1 + \ldots + Y_r) = r \cdot \mathbb{E}(Y_1) = \frac{r}{p}$$

Definition 15.5. *Let X be a discrete random variable. The* **variance** *of X, denoted by $\text{Var}(X)$, is defined as*

$$\text{Var}(X) = \mathbb{E}([X - \mathbb{E}(X)]^2)$$

Furthermore, the **standard deviation** *of X, denoted by σ_X, is defined as*

$$\sigma_X = \sqrt{\text{Var}(X)}$$

Remark 15.6. *Variance and standard deviation indicate how far or how close X is distributed around its average.*

Lemma 15.7. *Let X be a discrete random variable. Then*

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$$

Proof. We have

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}([X - \mathbb{E}(X)]^2) = \mathbb{E}([X^2 - 2X\mathbb{E}(X) + [\mathbb{E}(X)]^2)\\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)\cdot\mathbb{E}(X) + \mathbb{E}([\mathbb{E}(X)]^2) = \mathbb{E}(X^2) - 2\mathbb{E}(X)^2 + [\mathbb{E}(X)]^2\\ &= \mathbb{E}(X^2) - [\mathbb{E}(X)]^2\end{aligned}$$

□

Example 15.8. *(1) We throw a die, that is, $\Omega = \{1, 2, 3, 4, 5, 6\}$ and let X denote the outcome. Then $\mathbb{E}(X) = 3.5$. Furthermore*

$$\mathbb{E}(X^2) = \sum_{k=1}^{6} k^2 \cdot P_X(k) = \frac{1}{6}\cdot(1 + 4 + 9 + 16 + 25 + 36) = \frac{91}{6} = 15.166...$$

Hence $\mathrm{Var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = 15.166 - (3.5)^2 = 2.91$. *Furthermore* $\sigma_X = \sqrt{2.91} = 1.708$.

(2) Let $X \sim \mathrm{Bernoulli}(p)$. Then $R_X = \{0, 1\}$ and $\mathbb{E}(X) = p$. Furthermore

$$\mathbb{E}(X^2) = 0^2 \cdot P_X(0) + 1^2 \cdot P_X(1) = 0 \cdot (1 - p) + 1 \cdot p = p$$

and so $\mathrm{Var}(X) = p - p^2 = p(1 - p)$.

Lemma 15.9. (1) *Let X be a random variable and $a, b \in \mathbb{R}$. Then*

$$\mathrm{Var}(a \cdot X + b) = a^2 \cdot \mathrm{Var}(X).$$

(2) *Let $X_1, \ldots, X_n$ be independent random variables. Then*

$$\mathrm{Var}(X_1 + \ldots + X_n) = \mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n).$$

Proof. (1) Set $Y = a \cdot X + b$. Then $\mathbb{E}(Y) = a \cdot \mathbb{E}(X) + b$. Now

$$\begin{aligned}\mathrm{Var}(Y) &= \mathbb{E}[(Y - \mathbb{E}(Y))^2] = \mathbb{E}[(a \cdot X + b - a \cdot \mathbb{E}(X) - b)^2]\\ &= \mathbb{E}[a^2 \cdot (X - \mathbb{E}(X))^2] = a^2 \cdot \mathbb{E}[(X - \mathbb{E}(X))^2] = a^2 \cdot \mathrm{Var}(X)\end{aligned}$$

□

Example 15.10. *Let $X \sim \mathrm{Binomial}(n, p)$. Then $X = \sum_{i=1}^{n} Y_i$, where $Y_i \sim \mathrm{Bernoulli}(p)$, for all $i = 1, \ldots, n$, and all Y_i are independent. Now*

$$\mathrm{Var}(X) = \mathrm{Var}(Y_1 + \ldots + Y_n) = n \cdot \mathrm{Var}(Y_1) = n \cdot p \cdot (1 - p)$$