W1&2

A is a proper subset of B A是B的真子集

$A \cap B$	The intersection of the sets A and B: $A \cap B = \{x \mid x \in A \; and \; x \in B\}$.
$A \setminus B$	The *difference* of the sets A and B: $A \setminus B = \{x \mid x \in A \; but \; x \notin B\}$.

Set	Description
The natural numbers, $\mathbb{N}$	The positive whole numbers $\{1, 2, 3, 4, \ldots\}$.
The integers, $\mathbb{Z}$	All positive and negative whole numbers $\{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$. Note that $\mathbb{N} \subset \mathbb{Z}$.
The rational numbers, $\mathbb{Q}$	These are fractions, such as $\frac{1}{2}, -\frac{5}{4}, 2$ (since $2 = \frac{2}{1}$). Since we can write every integer as a fraction, $\mathbb{Z} \subset \mathbb{Q}$.
The real numbers, $\mathbb{R}$	This set includes all of the rational numbers, as well as other numbers that can't be represented as fractions, such as $\sqrt{2}$ and π. The best way to think of it is by forming a number line where you fill in all the gaps between rational numbers. We have $\mathbb{Q} \subset \mathbb{R}$.
The complex numbers, $\mathbb{C}$	An extension of the real numbers to incorporate $\sqrt{-1}$ and similar numbers. We won't focus on these in this course.

rational number Q 有理数
real number R 实数
complex number C 虚数
0不是自然数，不属于N

The statement 'if A is true then B is true' can also be read as 'A implies B', and is often written as $A \Rightarrow B$.

If we have both $A \Rightarrow B$ and $B \Rightarrow A$, then we say 'A is true if and only if B is true'. This is often written as 'A is true iff B is true', where 'iff' is an abbreviation for 'if and only if'. This can also be written as $A \Leftrightarrow B$.

In this course (and in much of mathematics, statistics, and data science) we most commonly work with logarithms in base e.

This means we generally use $\log(x)$ to denote the *natural logarithm* of x. You may previously have seen this written instead as $\ln(x)$.

We will use $\log_k(x)$ to denote the logarithm of x in some base k, e.g., $\log_{10}(x)$ for log in base 10, $\log_2(x)$ for log in base 2.

Polynomials 多项式 quadratic 二次的
rational function 有理函数
piecewise function 分段函数
Logistic function

Logistic function

This is the type of function being plotted in Figures 1.14 and 1.15 on Week 1: Functions — Motivating Example: A Statistical Model. The logistic function $f(x) = \text{logistic}(x)$ is defined by:

$$\text{logistic}(x) = \frac{1}{1+e^{-(ax+b)}} = \frac{e^{ax+b}}{1+e^{ax+b}}$$

for constants $a, b \in \mathbb{R}$. This function is essential in *logistic regression*.

logit function

Logit function

Very much connected to the logistic function is the *logit* function, in that they are inverse functions of each other (as we will discover in the next sections). It is defined as:

$$\text{logit}(x) = \log(\frac{x}{1-x})$$

and is the inverse for the logistic function if $a = 1$ and $b = 0$ (see later in this section for what an inverse function is).

ReLU and softplus

ReLU and softplus

The rectified linear unit (ReLU) function is a common activation function in neural networks. It is defined as:

$$\text{ReLU}(x) = \begin{cases} x & \text{if } x \geq 0, \\ 0 & \text{if } x < 0 \end{cases}$$

The domain and range are the same as for $|x|$.

A smooth approximation to this function is the *softplus* function:

$$\text{softplus}(x) = \log(1+e^x)$$

the domain of function 函数的自变量范围
the range of function 函数的值范围

quotient 商

$$\text{the sum } f+g: \quad (f+g)(x) = f(x) + g(x)$$
$$\text{difference } f-g: \quad (f-g)(x) = f(x) - g(x)$$
$$\text{product } f \cdot g: \quad (f \cdot g)(x) = f(x)g(x)$$
$$\text{quotient } \frac{f}{g}: \quad \left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}$$

复合函数

Definition: Composition of functions

$(f \circ g)(x) = f(g(x))$ and $(g \circ f)(x) = g(f(x))$

Domains of $f \circ g$ and $g \circ f$

In general, the domain of $f \circ g$ (or $g \circ f$) depends on both the domain of f and g. The composition $f \circ g$ has domain $\{x \in \mathcal{D}(g) \mid g(x) \in \mathcal{D}(f)\}$.

horizontal 水平的

The inverse function of $f(x)$ is usually written as $f^{-1}(x)$,

linear operator.

Definition: Linear operator

An operator L is *linear* if for all functions f and g, and every scalar $c \in \mathbb{R}$,

$$\begin{aligned} L(cf) &= cL(f) \\ L(f+g) &= L(f) + L(g) \end{aligned}.$$

1. $\sum_{i=1}^{n} 1 = \underbrace{1 + \cdots + 1}_{n} = n$
2. $\sum_{j=0}^{n} ar^j = a + ar + ar^2 + \cdots + ar^n = a\frac{1-r^{n+1}}{1-r}$ is the geometric sum.
3. $\sum_{i=1}^{n} i = 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$
4. $\sum_{i=1}^{n} i^2 = \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6} = \frac{n}{6}(2n^2 + 3n + 1) = \frac{n}{6}(2n+1)(n+1)$

Multiple Summation

\\(

$$\sum_{1\le(j,k)\le3} a_j b_k = a_1b_1 + a_1b_2 + a_1b_3 +$$

$$a_2b_1 + a_2b_2 + a_2b_3 +$$
$$a_3b_1 + a_3b_2 + a_3b_3$$

\\)

$$\sum_{j=1}^{3}\left(\sum_{k=1}^{3} a_j b_k\right)$$

结合律　Associativity

$$\sum_{j\,\in\,J}\ \sum_{k\,\in\,K} a_{j,k} = \sum_{k\,\in\,K}\sum_{j\,\in\,J} a_{j,k}$$

分配律　Distributivity

$$\sum_{j\in J,k\in K} a_j b_k = \left(\sum_{j\in J} a_j\right)\left(\sum_{k\in K} b_k\right)$$

$$\sum_{j=1}^{3}\left(\sum_{k=1}^{3} a_j b_k\right) = \left(\sum_{j=1}^{3} a_j\right)\left(\sum_{k=1}^{3} b_k\right)$$

判断级数是否收敛

p-series

The p-series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ is convergent if $p > 1$ and divergent if $p \leq 1$.

Geometric series

The geometric series $\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \ldots$ is convergent if $|x| < 1$ and divergent if $|x| \geq 1$

Vanishing criterion

If a series $\sum_{n=0}^{\infty} a_n$ converges then $a_n \to 0$ as $n \to \infty$. Another way of saying this is that if $\lim_{n\to\infty} a_n$ is **not** equal to zero, then the series $\sum_{n=0}^{\infty} a_n$ diverges.

vanishing criterion 反之不成立

$\lim_{n\to\infty} a_n = 0$ does $\underline{\underline{\text{not}}}$ imply $\sum_{n=0}^{\infty} a_n$ converges !!

$\sum_{n=1}^{\infty} \frac{1}{n} \longleftarrow$ DIVERGES!! (even though $\lim_{n\to\infty} \frac{1}{n} = 0$)