Masd Notes

Author: Rasmus Ladefoged

Department of Computer Science
University of Copenhagen
Denmark
November 6, 2024

Contents

(a) $\lim_{x\to -3^-} h(x)$ (b) $\lim_{x\to -3^+} h(x)$ (c) $\lim_{x\to -3} h(x)$
(d) $h(-3)$ (e) $\lim_{x\to -0^-} h(x)$ (f) $\lim_{x\to 0^+} h(x)$
(g) $\lim_{x\to 0} h(x)$ (h) $h(0)$ (i) $\lim_{x\to 2} h(x)$
(j) $h(2)$ (k) $\lim_{x\to 5^+} h(x)$ (l) $\lim_{x\to 5^-} h(x)$

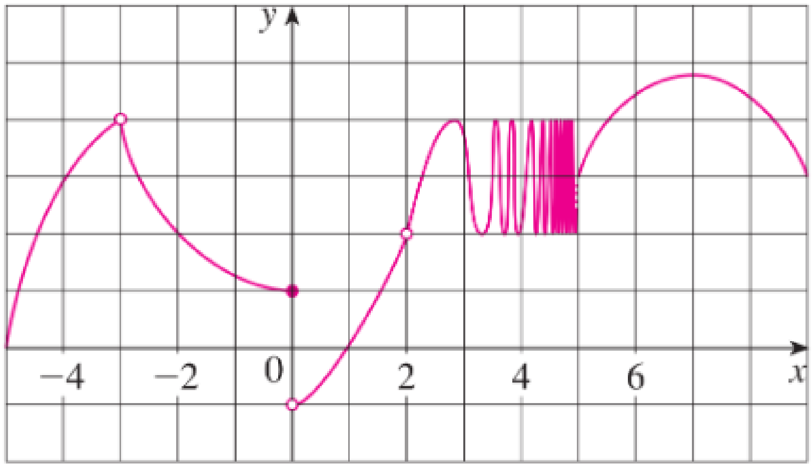

Figure 1: Heaviside function

1 Epsilon Funtion

En epsilon funktion er en funktion som er 0 i nul $fn(0) = 0$
For ethhvert heltal $\epsilon > 0$, findes der et heltal $f\epsilon$ sådan så snart IxI $< \frac{1}{k}$
der findes kun en epsilon funktion

$$f(x) = \delta(x_0) + \epsilon(x - x_0)$$

Den her funktion kigger på differencen mellem x og x_0, som bliver mindre og mindre jo nærmere man kommer på limit

1.1 Heaviside step function

Is where you go from either the left or right side of the point towards the limit, it looks like this.
The way you solve this is by looking at the limit sign value. For a) it would be:

$$limx- > -3^- h(x)$$

Here lim is the limit we are approaching, we are moving from x to -3, here x is usually a neighbouring value, either -2 or -3. The way you find which of these two it is, is by looking at the sign value of the limit function, which is -. This can be found -3^-. There are 3 signs in total

- - go from left
- $+$ go from right
- Is there a value at the point

So to solve this you just return the y value that can be traced from the direction to the given point. If there is no continous direction or an inconclusive one, then return None.
If there is an empty circle on a point, that means it is undefined at that position, therefore no value is given
The answers are:

- a) $= 4$

- b) = 4
- c) = 4
- d) = None, here the point is undefined and we are not approaching it, therefore there is no value on this exact point
- e) = 1, as we approach from the left due to the limit signage
- f) = -1 as we approach from the right due to the limit signage
- g) = None, here there are two options and we are not given a direction to approach from so it's undecidable
- h) = None, there are two values here and one of them is not defined
- i) = 2, as we are approaching it, but at 2 exactly there is no value
- j) = None there is no value directly on 2
- k) None, the line oscillates to much going left to right and therefore a point and direction cannot be determined
- l) = 3, there is no oscilliation from right to left and therefore a point can be determined

2 Algebraic Rules for Limits of Sequences

Below, we state some classical rules for determining the limits of sequences without proof. Let $(a_n)_n$ and $(b_n)_n$ be sequences of real numbers.

1. Assume $\lim_{n\to\infty} a_n = a$, where a could be $\pm\infty$, and let $k \in \mathbb{R}$ be any real number. Then:
$$\lim_{n\to\infty} ka_n = ka = k \lim_{n\to\infty} a_n.$$
If $\lim_{n\to\infty} a_n = \pm\infty$, then $\lim_{n\to\infty} ka_n = \text{sign}(k) \lim_{n\to\infty} a_n$.
2. If $(a_n)_n$ and $(b_n)_n$ converge to a and b (finite), then:
$$\lim_{n\to\infty} (a_n + b_n) = a + b = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n.$$
If $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = +\infty$, then the limit of the sum is $+\infty$. Similarly, if both limits are $-\infty$, the sum will tend to $-\infty$. If one sequence converges to $+\infty$ and the other to $-\infty$, the result is indeterminate and requires further computation.
3. If $(a_n)_n$ and $(b_n)_n$ converge to a and b (finite), then:
$$\lim_{n\to\infty} (a_n b_n) = ab = \lim_{n\to\infty} a_n \lim_{n\to\infty} b_n.$$
If both a_n and b_n converge to $\pm\infty$, the product will tend to $\pm\infty$ depending on the signs. If one sequence converges to $\pm\infty$ and the other to 0, the result is indeterminate and further computation is necessary.
4. If $\lim_{n\to\infty} a_n = a \neq 0$ (finite), then:
$$\lim_{n\to\infty} \frac{1}{a_n} = \frac{1}{a}.$$
If $\lim_{n\to\infty} a_n = \pm\infty$, then $\lim_{n\to\infty} \frac{1}{a_n} = 0$.
5. More generally, if $(a_n)_n$ and $(b_n)_n$ converge to a and b (finite) with $b \neq 0$, then:
$$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}.$$
If $a \neq 0$ and $\lim_{n\to\infty} b_n = \pm\infty$, then $\lim_{n\to\infty} \frac{a_n}{b_n} = 0$. If $a = 0$ and $\lim_{n\to\infty} b_n = \pm\infty$, the limit is indeterminate.
6. If $\lim_{n\to\infty} a_n = a$ and f is a continuous function at a, such that $f(a_n)$ is defined for all a_n (or at least for sufficiently large n), then:
$$\lim_{n\to\infty} f(a_n) = f(a).$$
This property is often used as the definition of a continuous function.

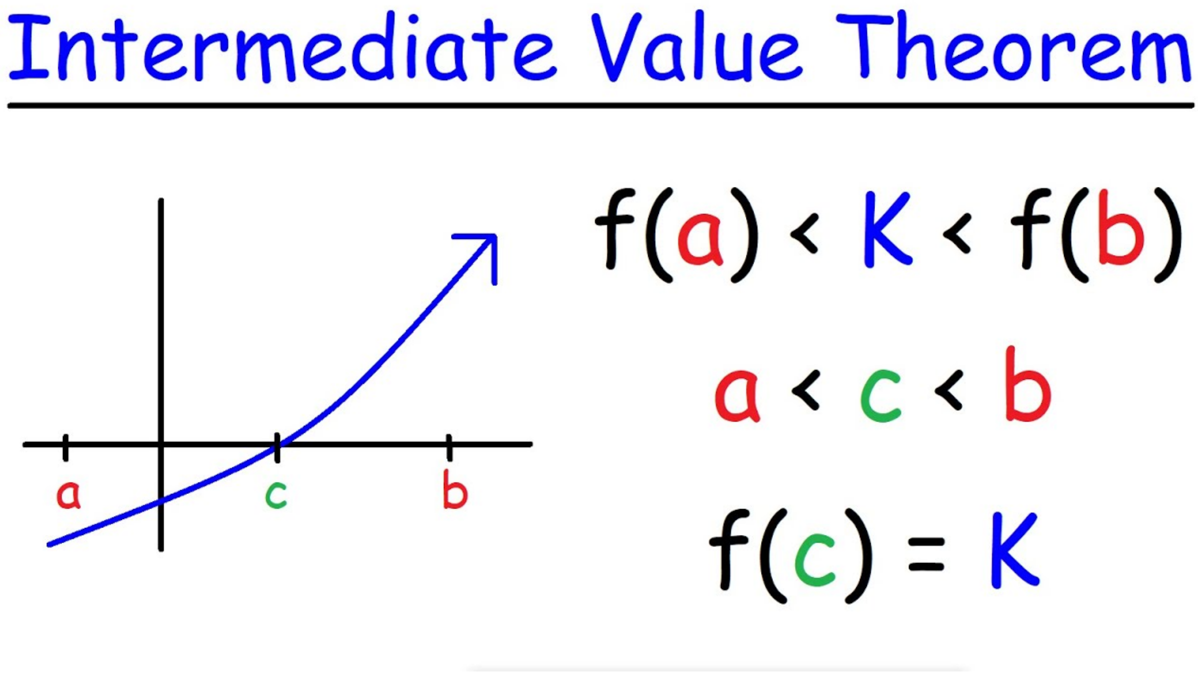

Figure 2: Intermediate Value Theorem

3 Intermediate values theorem

The intermediate values theorem is the idea that moving from a negative a value to a positive b value in a continuous line, must mean that we cross c = 0 at some point. Here C is called the root. The formal definition is found below:

3.1 Theorem 4.8 (Intermediate Values Theorem)

The Intermediate Value Theorem states that if a function $f : [a, b] \to \mathbb{R}$ is continuous on a closed interval $[a, b]$, and if $f(a)$ and $f(b)$ have opposite signs (i.e., $f(a) \cdot f(b) < 0$), then there exists at least one c in the interval $[a, b]$ such that $f(c) = 0$. Here, c is called the root of the function within the interval.
There are differnet ways to find C:

3.2 Bisection Method

Take the middle most value between a and b and plot it as d:

$$d = \frac{a + b}{2}$$

Next we check if f(d) = 0, if yes, then c = d and we have solved the problem. If not we check if d has the same sign as a. If yes then the new interval is [d, b] if not then it is [a, d]. We then repeat these two steps until f(d) = 0.

The reason why we check for the sign is that if both ranges of the interval have the same sign, then we have already crossed 0 and therefore missed the c value.

Formal description below:

3.3 Bisection Definition

The bisection algorithm searches for a vanishing point of a continuous function f on a range $[a, b]$ with $f(a)f(b) < 0$, i.e., the sign of $f(a)$ and the one of $f(b)$ differ. By the intermediate value theorem, there is at least one point y in $[a, b]$ with $f(y) = 0$. The bisection algorithm proceeds by repeatedly cutting $[a, b]$ in two and deciding which semi-interval to use according to function signs. It proceeds as follows:

- **Data:** $f : [a, b] \to \mathbb{R}$ with $f(a)f(b) < 0$, $\epsilon > 0$ (precision).
- **Initialization:** $a_0 = a$, $b_0 = b$.
- **For** $n = 0$ **until convergence:**
  1. Compute $c_n = \frac{a_n+b_n}{2}$ (midpoint between a_n and b_n).
  2. If $|f(c_n)| < \epsilon$, we have converged, return c_n.
  3. Else if $f(a_n)f(c_n) < 0$: $a_{n+1} = a_n$, $b_{n+1} = c_n$.
  4. Else $a_{n+1} = c_n$, $b_{n+1} = b_n$ (because we should have $f(c_n)f(b_n) < 0$ in that case).

3.4 Bisection Algorithm Example

Exercise 3 (Bisection method). The bisection algorithm is described in [?], where it is used to find a zero (root) of the equation $f(x) = 0$ with

$$f(x) = \frac{13x}{99} + \frac{x^2}{99} - \frac{269}{99} - \frac{47}{33}, \quad x \in [-5, 6].$$

- First compute $f(-5)$ and $f(5)$. Explain why there is a root in that range.

First compute f (5) and f (5). Explain why there is a root in that range:

$$f(x) = \frac{13}{99}x^3 - \frac{x^2}{99} + \frac{269}{99}x - \frac{47}{33} \tag{3.1}$$

We will start by computing f(-5), which is left limit of the range x $\in [-5, 5]$:

$$f(a) = f(-5) = \frac{13}{99} \cdot -5^3 - \frac{-5^2}{99} + \frac{269}{99} \cdot -5 - \frac{47}{33} \tag{3.2}$$

$$f(a) = f(-5) = -\frac{3086}{99} \tag{3.3}$$

Next we do f(5):

$$f(b) = f(5) = \frac{13}{99} \cdot 5^3 - \frac{5^2}{99} + \frac{269}{99} \cdot 5 - \frac{47}{33} \tag{3.4}$$

$$f(b) = f(5) = \frac{2804}{99} \tag{3.5}$$

As can be seen we have two results with opposite signs when we compute the limits of the range. This means that according to the Intermediate Values Theorem there must be a root point between the two where it crosses y = 0.
Next we will compute the middle point between the two to approach the limit:

$$c = \frac{-5 + 5}{2} = 0 \tag{3.6}$$

$$f(c) = f(0) = \frac{13}{99} \cdot 0^3 - \frac{0^2}{99} + \frac{269}{99} \cdot 0 - \frac{47}{33} \tag{3.7}$$

$$f(c) = f(0) = -\frac{141}{99} \tag{3.8}$$

We now see that f(a) and f(c) shares their sign, therefore we now know that a root lies in the range [c, b]. Therefore we will now compute the point between f(c) and f(b):

$$f(a) = f(c) \tag{3.9}$$

$$f(b) = f(b) \tag{3.10}$$

$$c = \frac{0 + 5}{2} = \frac{5}{2} \tag{3.11}$$

$$f(c) = f(\frac{5}{2}) = \frac{13}{99} \cdot \frac{5^3}{2} - \frac{\frac{5^2}{2}}{99} + \frac{269}{99} \cdot \frac{5}{2} - \frac{47}{33} \tag{3.12}$$

$$f(c) = f(\frac{5}{2}) = \frac{5827}{792} \tag{3.13}$$

f(c) shares sign with f(b), therefore we now look at the range [a, c].

4 Writing Proofs

Direct Proof: Shows that if p is true, then q follows directly by applying logical deductions and known facts.

Proof by Contradiction: Assumes the negation of the statement and shows that this assumption leads to a contradict

Proof by Induction:

- **Base Case:** Show the statement is true for the initial value.
- **Inductive Step:** Assume the statement is true for n, then prove it is true for n+1.

Proof by Contrapositive: Shows that if $\neg q$ is true, then $\neg p$ must be true, thus proving $p \implies q$.

Proof by Exhaustion: Divides the problem into a finite number of cases and proves each case individually.

4.1 Induction proof example

1. Show by induction on $n \in \mathbb{N}$ that

$$\sum_{i=1}^{n} i^3 - i^2 = \frac{(n-1)n(n+1)(3n+2)}{12}. \tag{4.1}$$

We start with the base case which is found by looking at the sum, which goes from i = 1 to n. Therefore the base case must be 1. We then solve both the left and right side with n = 1:

$$\sum_{i=1}^{1} i^3 - i^2 = \frac{(1-1) \cdot 1 \cdot (1+1)(3 \cdot 1 + 2)}{12} \tag{4.2}$$

$$1^3 - 1^2 = \frac{1 - 1 \cdot (1+1)(3+2)}{12} \tag{4.3}$$

$$1^3 - 1^2 = \frac{0 \cdot (1+1)(3+2)}{12} \tag{4.4}$$

$$1^3 - 1^2 = \frac{0}{12} \tag{4.5}$$

$$0 = 0 \tag{4.6}$$

This is the base case solved, next we do the induction step where we set n = n + 1.

$$\sum_{i=1}^{n+1} i^3 - i^2 = \frac{((n+1)-1)(n+1)((n+1)+1)(3(n+1)+2)}{12}. \tag{4.7}$$

We know that in the range n the left side equals

$$\frac{(n-1)n(n+1)(3n+2)}{12} \tag{4.8}$$

Therefore with n = n + 1 it would be:

$$\frac{(n-1)n(n+1)(3n+2)}{12} + (n+1)^3 - (n+1)^2 \tag{4.9}$$

Here we have inserted n+1 as the final i value, next we reduce this expression:

$$\frac{(n^2-n)(n+1)(3n+2)}{12} + (n+1)^3 - (n+1)^2 \tag{4.10}$$

$$\frac{(n^3-n)(3n+2)}{12} + (n+1)^3 - (n+1)^2 \tag{4.11}$$

$$\frac{3n^4+2n^3-3n^2-2n}{12} + (n+1)^3 - (n+1)^2 \tag{4.12}$$

$$\frac{3n^4+2n^3-3n^2-2n}{12} + (n+1)(n+1)(n+1) - (n+1)(n+1) \tag{4.13}$$

$$\frac{3n^4+2n^3-3n^2-2n}{12} + (n^2+2n+1)(n+1) - (n^2+2n+1) \tag{4.14}$$

$$\frac{3n^4+2n^3-3n^2-2n}{12} + (n^3+2n^2+3n+1) - (n^2+2n+1) \tag{4.15}$$

$$\frac{3n^4+2n^3-3n^2-2n}{12} + n^3+2n^2+n \tag{4.16}$$

$$\frac{3n^4+2n^3-3n^2-2n}{12} + \frac{12n^3+24n^2+12n}{12} \tag{4.17}$$

$$\frac{3n^4+2n^3-3n^2-2n+12n^3+24n^2+12n}{12} \tag{4.18}$$

$$\frac{3n^4+14n^3+21n^2+10n}{12} \tag{4.19}$$

We have now solved the left side, we now need to do the same for the right side, but since this is the closed-form expression, we can simply insert n + 1 directly into the formula:

$$\frac{((n+1)-1)(n+1)((n+1)+1)(3(n+1)+2)}{12} \tag{4.20}$$

$$\frac{n(n+1)(n+2)(3n+5)}{12} \tag{4.21}$$

$$\frac{(n^2+n)(n+2)(3n+5)}{12} \tag{4.22}$$

$$\frac{(n^3+3n^2+2n)(3n+5)}{12} \tag{4.23}$$

$$\frac{3n^4+5n^3+9n^3+15n^2+6n^2+10n}{12} \tag{4.24}$$

$$\frac{3n^4+14n^3+21n^2+10n}{12} \tag{4.25}$$

Now we compare the two sides and see if they match, if they do then we have proven the function via induction:

$$\frac{3n^4+14n^3+21n^2+10n}{12} = \frac{3n^4+14n^3+21n^2+10n}{12} \tag{4.26}$$

4.2 Induction Proof Example 2

Show by induction that $f(x) = x^n$ is continuous for any $n \geq 0$
To prove the inductive hypothesis: x^n is continuous, we start with the base case of n = 0:

$$f(x) = x^0 = 1 \tag{4.27}$$

This is a constant function and therefore continuous. Which proves the base case.
Next we do the induction step where n = n+1:

$$x^{n+1} = x^{n+1} = x^n \cdot x \tag{4.28}$$

The product of two continuous functions must also be continuous. As can be seen then we are simply taking the product of $x \cdot x$ n times and as the base case was continuous, which we proved with base case. Since the inductive step was also continuous, then we can conclude based on induction that the function $f(x) = x^n$ is continuous for any $n \geq 0$

5 Derivative

The derivative of a function $f(x)$ at a point x is defined as:

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

where h is an infinitesimally small increment in x.
Two different ways of writing it, f'(x) or $\frac{d}{dx}$

5.1 Common Derivatives

f	f'	f	f'
x^n	nx^{n-1}	$\ln x$	$\frac{1}{x}$
e^x	e^x	a^x	$\ln a \cdot a^x$
e^{ax}	ae^{ax}	$\sin x$	$\cos x$
$\cos x$	$-\sin x$	$-\sin x$	$-\cos x$
$-\cos x$	$\sin x$	$\tan x$	$\sec^2 x$
$\sqrt{x}$	$\frac{1}{2\sqrt{x}}$	$\arctan x$	$\frac{1}{1+x^2}$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$	$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$
f^n	$nf'f^{n-1}$	$\frac{1}{f}$	$-\frac{f'}{f^2}$
e^f	$f'e^f$	$\ln f$	$\frac{f'}{f}$
$\sinh x$	$\cosh x$	$\cosh x$	$\sinh x$
$\tanh x$	$\operatorname{sech}^2 x$	$\coth x$	$-\operatorname{csch}^2 x$
$\sec x$	$\sec x \tan x$	$\log_b x$	$\frac{1}{x \ln b}$

5.2 Power Rule

5.2.1 Power Rule example

$$f(x) = x^5 f'(x) = 5x^4 \tag{5.1}$$

$$f(x) = 10x f'(x) = 10 * 1 = 10 \tag{5.2}$$

$$f(x) = x^2 * 5x^4 \text{reduce the expression} f(x) = 6x^6 f'(x) = 30x^5 \tag{5.3}$$

The **Power Rule** for derivatives states that if $f(x) = x^n$ (where n is a constant), then:

$$f'(x) = nx^{n-1}$$

5.3 Quotient Rule

The **Quotient Rule** for derivatives is given by:

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$$

where f' and g' are the derivatives of f and g respectively, and $g \neq 0$.

5.3.1 Quotient Rule Examples

$$\frac{d}{dx}\left(\frac{\ln x}{x^4}\right) \tag{5.4}$$

To solve this we need to use the quotient rule:

Derivative Rules

1. Constant Rule : $\frac{d}{dx}(c) = 0$

2. Constant Multiple Rule : $\frac{d}{dx}[cf(x)] = cf'(x)$

3. Power Rule : $\frac{d}{dx}(x^n) = nx^{n-1}$

4. Sum Rule : $\frac{d}{dx}[f(x) + g(x)] = f'(x) + g'(x)$

5. Difference Rule : $\frac{d}{dx}[f(x) - g(x)] = f'(x) - g'(x)$

6. Product Rule : $\frac{d}{dx}[f(x)g(x)] = f(x)g'(x) + g(x)f'(x)$

7. Quotient Rule : $\frac{d}{dx}\left[\frac{f(x)}{g(x)}\right] = \frac{g(x)f'(x) - f(x)g'(x)}{[g(x)]^2}$

8. Chain Rule : $\frac{d}{dx}f[g(x)] = f'[g(x)]g'(x)$

Figure 3: Derivative Rules

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2} \tag{5.5}$$
$$f(x) = \ln x \tag{5.6}$$
$$f'(x) = \frac{1}{x} \tag{5.7}$$
$$g(x) = x^4 \tag{5.8}$$
$$g'(x) = 4x^3 \tag{5.9}$$

We can now insert it into the quotient function

$$\frac{d}{dx} = \frac{\frac{1}{x} \cdot x^4 - \ln(x) \cdot 4x^3}{(x^4)^2} \tag{5.10}$$
$$\tag{5.11}$$

We can now reduce the function:

$$\frac{d}{dx} = \frac{\frac{x^4}{x} - \ln(x) \cdot 4x^3}{x^8} \tag{5.12}$$
$$\frac{d}{dx} = \frac{x^3 - \ln(x) \cdot 4x^3}{x^8} \tag{5.13}$$
$$\frac{d}{dx} = \frac{x^3(1 - 4\ln(x))}{x^8} \tag{5.14}$$
$$\frac{d}{dx} = \frac{1 - 4\ln(x)}{x^5} \tag{5.15}$$
$$\tag{5.16}$$

We have now solved the derivative

5.4 Product Rule

The **Product Rule** states that the derivative of the product of two functions $h(x)$ and $g(x)$ is:

$$(hg)' = h'g + hg'$$

where h' and g' are the derivatives of h and g respectively.

5.4.1 Product Rule example

$$f(x) = x^2 * 5x^4 \tag{5.17}$$
$$f'(x) = h'g + hg' \tag{5.18}$$
$$h = x^2 \tag{5.19}$$
$$h' = 2x \tag{5.20}$$
$$g = 5x^4 \tag{5.21}$$
$$g' = 5 * 4x^3 = 20x^3 \tag{5.22}$$
$$f'(x) = 2x * 5x^4 + x^2 * 20x^3 \tag{5.23}$$
$$f'(x) = 10x^5 + 20x^5 = 30x^5 \tag{5.24}$$
$$\tag{5.25}$$

Using power rule here is simpler

$$f(x) = 6 * x^3 \tag{5.26}$$
$$f'(x) = h'g + hg' \tag{5.27}$$
$$h = 6 \tag{5.28}$$
$$h' = 0 \tag{5.29}$$
$$g = x^3 \tag{5.30}$$
$$g' = 3x^2 \tag{5.31}$$
$$f'(x) = 0 * x^3 + 6 * 3x^2 \tag{5.32}$$
$$f'(x) = 18x^2 \tag{5.33}$$
$$\tag{5.34}$$

5.5 Chain Rule

The **Chain Rule** states that the derivative of a composite function $f(g(x))$ is:

$$\frac{d}{dx}[f(g(x))] = f'(g(x)) \cdot g'(x)$$

where f' is the derivative of f with respect to g, and g' is the derivative of g with respect to x.

5.5.1 Chain Rule examples

$\frac{d}{dx}\sqrt{2x^2 + 2x + 4}$ To solve this we need to use the chain rule:

$$f(u) = \sqrt{2x^2 + 2x + 4} \tag{5.35}$$
$$f'(u) = \frac{1}{2\sqrt{2x^2 + 2x + 4}} \cdot \frac{d}{du}uu = 2x^2 + 2x + 4u' = 4x + 2 \tag{5.36}$$

We now insert u':

$$f'(u) = \frac{1}{2\sqrt{x^2 + 2x + 4}} \cdot (4x + 2)f'(u) = \frac{4x + 2}{2\sqrt{x^2 + 2x + 4}} \tag{5.37}$$

5.6 Complicated examples

5.6.1 $\frac{d}{dx}\sin(x)e^{-\cos(x)^2}$

To solve this we need to start with the product rule:

$$\frac{d}{dx}\left(f(x) \cdot g(x)\right) = f'(x) \cdot g(x) + f(x) \cdot g'(x) \tag{5.38}$$
$$f(x) = \sin(x) \tag{5.39}$$
$$f'(x) = \cos(x) \tag{5.40}$$
$$g(x) = e^{-\cos(x)^2} \tag{5.41}$$
$$\tag{5.42}$$

Now to solve g'(x) we need to use the chain rule. We will rewrite it to g(u) to signify the inner function:

$$g(u) = e^{-\cos(x)^2} \tag{5.43}$$
$$g'(u) = e^{-\cos(x)^2}\frac{d}{du}u \tag{5.44}$$
$$u = -\cos(x)^2 \tag{5.45}$$

To solve u' we need to use the chain rule again:

$$u(t) = -\cos(x)^2 \tag{5.46}$$
$$u'(t) = 2 - \cos(x) \cdot \frac{d}{dt}t \tag{5.47}$$
$$t = \cos(x) \tag{5.48}$$
$$t' = -\sin(x) \tag{5.49}$$
$$\tag{5.50}$$

We now finish u'(t):

$$u'(t) = 2 - \cos(x) \cdot - \sin(x) \tag{5.51}$$
$$u'(t) = 2\cos(x) \cdot \sin(x) \tag{5.52}$$

Now we can complete g'(u):

$$g'(u) = e^{-\cos(x)^2} \cdot 2\cos(x) \cdot \sin(x) \tag{5.53}$$
$$\tag{5.54}$$

We can now complete the initial function using the product rule:

$$\frac{d}{dx}\left(f(x) \cdot g(x)\right) = f'(x) \cdot g(x) + f(x) \cdot g'(x) \tag{5.55}$$
$$\frac{d}{dx}\left(\cos(x) \cdot e^{-\cos(x)^2}\right) = \cos(x) \cdot e^{-\cos(x)^2} + \sin(x) \cdot e^{-\cos(x)^2} \cdot 2\cos(x)\sin(x) \tag{5.56}$$

We have now solved the derivative

5.6.2 $\frac{d}{dx}\left(\frac{\ln(1+e^{qx})}{q}\right)$

It is important to note here that q is a constant, and the variable is x, as denoted by $\frac{d}{dx}$. We will therefore start by factoring out $\frac{1}{q}$:

$$\frac{1}{q} \cdot \frac{d}{dx}\left(\ln(1 + e^{qx})\right) \tag{5.57}$$

Now, we use the chain rule on $\ln(1 + e^{qx})$:

$$f(u) = \ln(1 + e^{qx}) \tag{5.58}$$
$$f'(u) = \frac{1}{1 + e^{qx}} \cdot \frac{d}{du}(1 + e^{qx}) \tag{5.59}$$
$$u = 1 + e^{qx} \tag{5.60}$$
$$u' = q \cdot e^{qx} \tag{5.61}$$

Now we solve for $f'(u)$:

$$\frac{d}{dx}\ln(1 + e^{qx}) = \frac{1}{1 + e^{qx}} \cdot q \cdot e^{qx} \tag{5.62}$$
$$= \frac{q \cdot e^{qx}}{1 + e^{qx}} \tag{5.63}$$

Now, inserting $\frac{1}{q}$ back into the function:

$$\frac{d}{dx}\left(\frac{\ln(1 + e^{qx})}{q}\right) = \frac{1}{q} \cdot \frac{q \cdot e^{qx}}{1 + e^{qx}} \tag{5.64}$$
$$= \frac{q \cdot e^{qx}}{q \cdot (1 + e^{qx})} \tag{5.65}$$

The q's now cancel out:

$$\frac{d}{dx}\left(\frac{\ln(1+e^{qx})}{q}\right) = \frac{e^{qx}}{1+e^{qx}} \tag{5.66}$$

5.7 Second Derivative

Second derivative is either written as f"(x) or $\frac{d^2}{dx^2}$. To solve the second derivative of something, you simply do the first derivative as we did above. Then we take the derivative of the solution we just found.

5.7.1 Second derivative examples

Basic example

$$f(x) = x^5 \tag{5.67}$$
$$f'(x) = 5x^4 \tag{5.68}$$
$$f''(x) = 4 \cdot 5x^3 = 20x^3 \tag{5.69}$$
$$\text{This can then be continued indefinitely} \tag{5.70}$$
$$f'''(x) = 3 \cdot 20x^2 = 60x^2 \tag{5.71}$$
$$f''''(x) = 2 \cdot 60x = 120x \tag{5.72}$$
$$f'''''(x) = 120 \cdot 1 = 120 \tag{5.73}$$
$$f''''''(x) = 0 \tag{5.74}$$
$$\tag{5.75}$$

The same principle applies no matter how many times you take the derivative of something.

Power rule example

$$f(x) = 6 * x^3 \tag{5.76}$$
$$f'(x) = h'g + hg' \tag{5.77}$$
$$h = 6 \tag{5.78}$$
$$h' = 0 \tag{5.79}$$
$$g = x^3 \tag{5.80}$$
$$g' = 3x^2 \tag{5.81}$$
$$f'(x) = 0 * x^3 + 6 * 3x^2 \tag{5.82}$$
$$f'(x) = 18x^2 \tag{5.83}$$
$$\tag{5.84}$$

To find the second derivative of $f(x) = 6 * x^3$, we simply take the solution and find the derivative of that:

$$f'(x) = 18x^2 \tag{5.85}$$
$$f''(x) = 2 \cdot 18x = 36x \tag{5.86}$$

More complicated example, using the complex quotient rule from earlier (5.65):

$$\frac{d^2}{dx^2}\left(\frac{\ln(1+e^{qx})}{q}\right) \tag{5.87}$$

We will start with the solution to the previous question since that's the first derivative:

$$\frac{d}{dx}\frac{e^{qx}}{1+e^{qx}} \tag{5.88}$$

To solve this, we will use the quotient rule:

$$\frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2} \tag{5.89}$$

$$f(x) = e^{qx} \tag{5.90}$$

$$f'(x) = e^{qx} \cdot q \tag{5.91}$$

$$g(x) = 1 + e^{qx} \tag{5.92}$$

$$g'(x) = e^{qx} \cdot q \tag{5.93}$$

Now, applying the quotient rule:

$$\frac{d^2}{dx^2}\left(\frac{e^{qx}}{1+e^{qx}}\right) = \frac{e^{qx} \cdot q \cdot (1+e^{qx}) - e^{qx} \cdot e^{qx} \cdot q}{(1+e^{qx})^2} \tag{5.94}$$

Finally we simplify the function:

$$= \frac{qe^{qx} \cdot (1+e^{qx}) - qe^{2qx}}{(1+e^{qx})^2} \tag{5.95}$$

$$= \frac{qe^{qx}}{(1+e^{qx})^2} \tag{5.96}$$

5.8 Partial Derivative

The partial derivative of a function $f(x, y, \ldots)$ with respect to the variable x is defined as:

$$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{f(x+h, y, \ldots) - f(x, y, \ldots)}{h}$$

where h is an infinitesimally small increment in x, and the other variables $y, \ldots$ are held constant.

5.8.1 Partial Derivative Explanation

Partial derivatives are used for functions that have two variables, x and y. You then need to split the differentiation in two. So you first differentiate on x and then you differentiate on y.

When you differentiate on x, then you treat y as a constant and when you differentiate on y you treat x as a constant.

5.8.2 Partial Derivative examples

$$f(x, y) = x^2y^4 - 5x^7 + 4y^8 \tag{5.97}$$

To start we differentiate on x:

$$\frac{\partial f}{\partial x} = 2x \cdot y^4 - 7 \cdot 5x^6 + 0 \tag{5.98}$$

$$\frac{\partial f}{\partial x} = 2x \cdot y^4 - 35x^6 \tag{5.99}$$

Here we see that y is treated as constant, therefore in the section x^2y^4, we simply treat it as $a = y^4$ $a \cdot x^2$. We then differentiate x but keep y as it is, since it's a constant. But in this section $5x^7 + 4y^8$, the y value is not multiplied with an x value and if therefore 0.
Next we do the y section:

$$\frac{\partial f}{\partial y} = x^2 \cdot 4 \cdot y^3 - 0 + 8 \cdot 4y^7 \tag{5.100}$$

$$\frac{\partial f}{\partial y} = 4y^3x^2 + 32y^7 \tag{5.101}$$

Here we do the same as before just inverse, treat y as the variable and x as a constant.

$$f(x, y) = e^{x^2y^3} \tag{5.102}$$

Here we need to use the chain rule first to solve the equation

$$\begin{align} f'(x, y) &= f'(u) \cdot u' \tag{5.103}\\ f(u) &= e^{x^2y^3} \tag{5.104}\\ f'(u) &= e^{x^2y^3} \tag{5.105}\\ u &= x^2y^3 \tag{5.106}\end{align}$$

Now to solve u' we need to do partial differentiation:

$$\begin{align} \frac{\partial u}{\partial x} &= 2xy^3 \tag{5.107}\\ \frac{\partial u}{\partial y} &= 3x^2y^2 \tag{5.108}\end{align}$$

We can now solve the two partial derivatives completely:

$$\begin{align} \frac{\partial f}{\partial x} &= e^{x^2y^3} \cdot 2xy^3 \tag{5.109}\\ \frac{\partial f}{\partial y} &= e^{x^2y^3} \cdot 3x^2y^2 \tag{5.110}\\ & \tag{5.111}\end{align}$$

This only works if f'(u) does not contain any of the partial derivatives. Otherwise you have to solve the partial derivative for each section

5.8.3 Complicated example

$$f(x, y) = \ln\left(\frac{x^2}{y}\right) \tag{5.112}$$

We start by defining u as:

$$u = \frac{x^2}{y} \tag{5.113}$$

To find the partial derivatives of f, we use the chain rule:
Partial Derivative with respect to x:

$$\begin{align} \frac{\partial u}{\partial x} &= \frac{2x}{y} \tag{5.114}\\ \frac{\partial f}{\partial x} &= \frac{1}{u} \cdot \frac{\partial u}{\partial x} = \frac{1}{\frac{x^2}{y}} \cdot \frac{2x}{y} \tag{5.115}\\ &= \frac{y}{x^2} \cdot \frac{2x}{y} \tag{5.116}\\ &= \frac{2x}{x^2} \tag{5.117}\\ &= \frac{2}{x} \tag{5.118}\end{align}$$

Partial Derivative with respect to y:

$$\frac{\partial u}{\partial y} = -\frac{x^2}{y^2} \tag{5.119}$$

$$\frac{\partial f}{\partial y} = \frac{1}{u} \cdot \frac{\partial u}{\partial y} = \frac{1}{\frac{x^2}{y}} \cdot \left(-\frac{x^2}{y^2}\right) \tag{5.120}$$

$$= \frac{y}{x^2} \cdot \left(-\frac{x^2}{y^2}\right) \tag{5.121}$$

$$= -\frac{1}{y} \tag{5.122}$$

5.9 Implicit Differentiation

In cases where y is defined implicitly by a relationship with x, rather than as an explicit function of x, we can use implicit differentiation to find $\frac{dy}{dx}$.
The process is as follows: 1. Differentiate both sides of the equation with respect to x, treating y as a function of x and applying the chain rule. 2. Collect terms containing $\frac{dy}{dx}$ and solve for $\frac{dy}{dx}$.
This technique is useful when isolating y is difficult.

5.9.1 Example 1: $x^2 + y^2 = 1$

We want to find $\frac{dy}{dx}$ for the equation $x^2 + y^2 = 1$, where y is defined implicitly as a function of x.
1. **Differentiate both sides with respect to x:**
The left side of the equation is $x^2 + y^2$. Differentiating each term with respect to x:

$$\frac{d}{dx}(x^2) + \frac{d}{dx}(y^2) = \frac{d}{dx}(1)$$

- The derivative of x^2 with respect to x is $2x$. - For y^2, we apply the chain rule. Since y is a function of x, we get:

$$\frac{d}{dx}(y^2) = 2y \cdot \frac{dy}{dx}$$

So, the differentiated equation becomes:

$$2x + 2y\frac{dy}{dx} = 0$$

2. **Solve for $\frac{dy}{dx}$:**
Now, isolate $\frac{dy}{dx}$ by moving terms involving x to the other side of the equation:

$$2y\frac{dy}{dx} = -2x$$

Divide both sides by $2y$:

$$\frac{dy}{dx} = -\frac{x}{y}$$

This gives the slope of y with respect to x, showing how y changes as x changes.

5.9.2 Example 2: $x^3 + y^3 = 6xy$

We want to find $\frac{dy}{dx}$ for the equation $x^3 + y^3 = 6xy$.
1. **Differentiate both sides with respect to x:**
Applying $\frac{d}{dx}$ to each term in $x^3 + y^3 = 6xy$:

$$\frac{d}{dx}(x^3) + \frac{d}{dx}(y^3) = \frac{d}{dx}(6xy)$$

- For x^3, the derivative is $3x^2$. - For y^3, use the chain rule, yielding $3y^2 \cdot \frac{dy}{dx}$. - For $6xy$, use the product rule: $\frac{d}{dx}(6xy) = 6y + 6x \cdot \frac{dy}{dx}$.
Substitute these results into the differentiated equation:

$$3x^2 + 3y^2\frac{dy}{dx} = 6y + 6x\frac{dy}{dx}$$

2. **Collect terms with $\frac{dy}{dx}$ on one side:**
Rearrange the equation to bring all terms with $\frac{dy}{dx}$ to one side:

$$3y^2\frac{dy}{dx} - 6x\frac{dy}{dx} = 6y - 3x^2$$

3. **Factor out $\frac{dy}{dx}$:**

$$\frac{dy}{dx}(3y^2 - 6x) = 6y - 3x^2$$

4. **Solve for $\frac{dy}{dx}$:**

$$\frac{dy}{dx} = \frac{6y - 3x^2}{3y^2 - 6x}$$

5.9.3 Example 3: $xy = 1$

We want to find $\frac{dy}{dx}$ for the equation $xy = 1$.
1. **Differentiate both sides with respect to x:**

$$\frac{d}{dx}(xy) = \frac{d}{dx}(1)$$

Applying the product rule on xy:

$$x\frac{dy}{dx} + y = 0$$

2. **Solve for $\frac{dy}{dx}$:**
Rearranging to isolate $\frac{dy}{dx}$:

$$x\frac{dy}{dx} = -y$$

$$\frac{dy}{dx} = -\frac{y}{x}$$

5.9.4 Example 4

Find $\frac{dy}{dx}$ given that $2x^3 + x^2y - xy^3 = 0$.
1. Differentiate both sides with respect to x:

$$\frac{d}{dx}(2x^3 + x^2y - xy^3) = \frac{d}{dx}(0)$$

$$6x^2 + 2xy + x^2\frac{dy}{dx} - y^3 - 3xy^2\frac{dy}{dx} = 0$$

2. Rearrange to isolate terms containing $\frac{dy}{dx}$:

$$(x^2 - 3xy^2)\frac{dy}{dx} = y^3 - 2xy - 6x^2$$

3. Solve for $\frac{dy}{dx}$:

$$\frac{dy}{dx} = \frac{y^3 - 2xy - 6x^2}{x^2 - 3xy^2}$$

5.9.5 Example 5

Inverse of example 4. Find $\frac{dx}{dy}$ given that $2x^3 + x^2y - xy^3 = 0$.
1. Differentiate both sides with respect to y:

$$\frac{d}{dy}(2x^3 + x^2y - xy^3) = \frac{d}{dy}(0)$$

2. Perform Differentiation

$$0 + x^2 - 3xy^2\frac{dx}{dy} + x^2\frac{dx}{dy} = 0$$

3. Rearrange to isolate terms containing $\frac{dx}{dy}$:

$$(x^2 - 3xy^2)\frac{dx}{dy} = -x^2$$

4. Solve for $\frac{dx}{dy}$:

$$\frac{dx}{dy} = \frac{-x^2}{x^2 - 3xy^2}$$
$$\frac{dx}{dy} = \frac{x^2}{3xy^2 - x^2}$$

This expression gives us $\frac{dx}{dy}$, illustrating how x changes with respect to changes in y, derived by treating x as a function of y and using implicit differentiation.

5.9.6 Example 6

Find $\frac{dy}{dx}$ if $x^3 + y^3 = 3xy$.
1. Differentiate both sides with respect to x:

$$\frac{d}{dx}(x^3 + y^3) = \frac{d}{dx}(3xy)$$

$$3x^2 + 3y^2\frac{dy}{dx} = 3y + 3x\frac{dy}{dx}$$

2. Rearrange to isolate $\frac{dy}{dx}$:

$$3y^2\frac{dy}{dx} - 3x\frac{dy}{dx} = 3y - 3x^2$$

$$(3y^2 - 3x)\frac{dy}{dx} = 3y - 3x^2$$

3. Solve for $\frac{dy}{dx}$:

$$\frac{dy}{dx} = \frac{3y - 3x^2}{3y^2 - 3x}$$
$$\frac{dy}{dx} = \frac{y - x^2}{y^2 - x}$$

6 Drawing Graphs

6.1 Drawing Graphs from a Function

To draw the graph of a function $f(x)$, it's helpful to analyze key features such as intercepts, asymptotes, and critical points. These characteristics provide a framework for sketching the graph accurately.

6.1.1 Steps to Draw the Graph of a Function

1. **Identify the domain** of $f(x)$: Determine all x-values for which the function is defined.
2. **Find intercepts**: Calculate $f(0)$ to find the y-intercept and solve $f(x) = 0$ for x-intercepts.
3. **Analyze asymptotes**: For rational functions, check for vertical asymptotes by identifying values of x that make the denominator zero, and find horizontal or oblique asymptotes by examining the end behavior as $x \to \pm\infty$.
4. **Compute the derivative** $f'(x)$: Find critical points by setting $f'(x) = 0$ and solve for x. These points indicate where the graph's slope is zero (potential peaks, valleys, or flat points).
5. **Determine concavity**: Compute the second derivative $f''(x)$ to analyze concavity. If $f''(x) > 0$, the function is concave up; if $f''(x) < 0$, it's concave down. Points where $f''(x) = 0$ may indicate inflection points.
6. **Plot points and sketch**: Using intercepts, critical points, and points of inflection, sketch the general shape of the graph.

6.1.2 Example 1

Given the function:

$$f(x) = \frac{x^2 - 1}{x - 2}$$

we will analyze its features and draw its graph.
1. **Domain**: The function is undefined at $x = 2$ (where the denominator is zero).
2. **Intercepts**: - y-intercept: Set $x = 0$:

$$f(0) = \frac{0^2 - 1}{0 - 2} = \frac{-1}{-2} = \frac{1}{2}$$

- x-intercepts: Set $f(x) = 0$:

$$\frac{x^2 - 1}{x - 2} = 0 \Rightarrow x^2 - 1 = 0 \Rightarrow x = \pm 1$$

3. **Asymptotes**: - Vertical asymptote at $x = 2$ (where $f(x)$ is undefined). - Horizontal asymptote: As $x \to \pm\infty$, the degree of the numerator and denominator are the same, so the asymptote is given by the ratio of the leading coefficients:

$$y = 1$$

4. **Derivative and Critical Points**: - Compute $f'(x)$ using the quotient rule:

$$f'(x) = \frac{(2x)(x - 2) - (x^2 - 1)(1)}{(x - 2)^2} = \frac{2x^2 - 4x - x^2 + 1}{(x - 2)^2} = \frac{x^2 - 4x + 1}{(x - 2)^2}$$

- Set $f'(x) = 0$ to find critical points:

$$x^2 - 4x + 1 = 0 \Rightarrow x = 2 \pm \sqrt{3}$$

5. **Concavity and Inflection Points**: - Compute the second derivative $f''(x)$ to determine concavity (omitted here for simplicity).
6. **Plotting**: - Plot the intercepts, asymptotes, and critical points. - Sketch the curve, using the behavior at intercepts and asymptotes as a guide.

6.1.3 Example 2

Consider the function:

$$g(x) = x^3 - 3x^2 + 4$$

we want to analyze this function to draw its graph.
1. **Domain**: The function $g(x)$ is defined for all $x \in \mathbb{R}$.
2. **Intercepts**: - y-intercept: Set $x = 0$:

$$g(0) = 0^3 - 3 \cdot 0^2 + 4 = 4$$

- x-intercepts: Set $g(x) = 0$:

$$x^3 - 3x^2 + 4 = 0$$

Solving this cubic equation for x gives the intercepts (which can be factored or solved numerically).
3. **Asymptotes**: - There are no asymptotes, as this is a polynomial function.
4. **Derivative and Critical Points**: - Compute $g'(x)$:

$$g'(x) = 3x^2 - 6x$$

- Set $g'(x) = 0$ to find critical points:

$$3x(x - 2) = 0$$

So, $x = 0$ and $x = 2$ are critical points.
5. **Concavity and Inflection Points**: - Compute $g''(x)$:

$$g''(x) = 6x - 6$$

- Set $g''(x) = 0$ to find inflection points:

$$6x - 6 = 0 \Rightarrow x = 1$$

6. **Plotting**: - Using the intercepts, critical points, and inflection points, we plot the function to visualize the shape of the graph.

7 Gradient and Hessian Matrix

- **Gradient**: The gradient of a function $f(x, y)$ is a vector of its first partial derivatives. It points in the direction of the steepest ascent of the function. The gradient is given by:

$$\nabla f(x,y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right)$$

- **Hessian Matrix**: The Hessian matrix of a function of two variables is a square matrix of second-order partial derivatives. It describes the local curvature of the function. The Hessian matrix is defined as:

$$H(f)(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

7.1 Examples

7.1.1 Example 1: Function $f(x, y) = x^2 + 3xy + y^2$

- **Gradient Calculation**:

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(x^2 + 3xy + y^2) = 2x + 3y$$
$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(x^2 + 3xy + y^2) = 3x + 2y$$
$$\nabla f(x,y) = (2x + 3y, 3x + 2y)$$

- **Hessian Matrix Calculation**:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}(2x + 3y) = 2$$
$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial y}(2x + 3y) = 3$$
$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial x}(3x + 2y) = 3$$
$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}(3x + 2y) = 2$$
$$H(f)(x,y) = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$$

7.1.2 Example 2: Function $f(x, y) = e^x \sin(y) + y^3$

- **Gradient Calculation**:

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}(e^x \sin(y) + y^3) = e^x \sin(y)$$
$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y}(e^x \sin(y) + y^3) = e^x \cos(y) + 3y^2$$
$$\nabla f(x,y) = \left(e^x \sin(y), e^x \cos(y) + 3y^2\right)$$

- **Hessian Matrix Calculation**:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}(e^x \sin(y)) = e^x \sin(y)$$
$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial y}(e^x \sin(y)) = e^x \cos(y)$$
$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial x}(e^x \cos(y)) = e^x \cos(y)$$
$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}(e^x \cos(y) + 3y^2) = -e^x \sin(y) + 6y$$
$$H(f)(x,y) = \begin{bmatrix} e^x \sin(y) & e^x \cos(y) \\ e^x \cos(y) & -e^x \sin(y) + 6y \end{bmatrix}$$

7.2 Finding Critical Points

Critical points occur where the gradient of the function is zero. This involves solving the system of equations obtained by setting each component of the gradient to zero.

7.2.1 General Steps

1. **Compute the Gradient:** Calculate the partial derivatives to form the gradient vector:

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right).$$

2. **Solve for Zero Gradient:** Set each component of the gradient to zero and solve the resulting system of equations:

$$\frac{\partial f}{\partial x} = 0, \quad \frac{\partial f}{\partial y} = 0.$$

   This yields the critical points (x, y).

7.3 Identifying the Nature of Critical Points

Once the critical points are found, the Hessian matrix is used to determine the nature of each point.

7.3.1 Analyze the Hessian

1. **Compute the Hessian Matrix:** Calculate the matrix of second-order partial derivatives:

$$H(f)(x, y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}.$$

2. **Evaluate the Hessian at Critical Points:** Substitute the critical points into the Hessian matrix.

3. **Determine the Nature of Critical Points:**
   - *Hessian < 0:* The Hessian is Indefinite, the function has a saddle point at that point.
   - *Hessian > 0:* The Hessian is Positive Definite, therefore it is either a local maximum or local minimum. To determine which, check the sign of f_{xx} ($\frac{\partial^2 f}{\partial x^2}$) (assuming f_{xx} matches the sign of other second derivatives):
     - $f_{xx} > 0$*:* the function has a local minimum at that point.
     - $f_{xx} < 0$*:* the function has a local maximum at that point.
   - *Hessian == 0:* The test is inconclusive; further analysis is required.

7.4 Example: Analyzing the Function $f(x, y) = x^2 - 2xy + 5y^2 + 3x - 4y$

7.4.1 Finding Critical Points

First, compute the gradient of f:

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) = (2x - 2y + 3, -2x + 10y - 4).$$

Set each component of the gradient to zero to find the critical points:

$$\begin{aligned} 2x - 2y + 3 &= 0, \\ -2x + 10y - 4 &= 0. \end{aligned}$$

Solving this system of equations:

$$\begin{aligned} 2x - 2y &= -3, \\ -2x + 10y &= 4. \end{aligned}$$

Add the two equations:

$$\begin{aligned} 2x - 2y - 2x + 10y &= -3 + 4, \\ 8y &= 1, \\ y &= \frac{1}{8}. \end{aligned}$$

Substitute $y = \frac{1}{8}$ back into the first equation:

$$\begin{aligned} 2x - 2\left(\frac{1}{8}\right) &= -3, \\ 2x - \frac{1}{4} &= -3, \\ 2x &= -3 + \frac{1}{4}, \\ 2x &= -\frac{11}{4}, \\ x &= -\frac{11}{8}. \end{aligned}$$

The critical point is $\left(-\frac{11}{8}, \frac{1}{8}\right)$.

7.5 Determining the Nature of the Critical Point

Compute the Hessian matrix of f:

$$H(f)(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & -2 \\ -2 & 10 \end{bmatrix}.$$

Evaluate the determinant of the Hessian at the critical point:

$$\begin{aligned} \det H(f) &= 2 \cdot 10 - (-2) \cdot (-2), \\ &= 20 - 4, \\ &= 16. \end{aligned}$$

Since the determinant is positive and the trace ($2 + 10 = 12$, all positive) is positive, the Hessian is positive definite, indicating that f has a local minimum at $\left(-\frac{11}{8}, \frac{1}{8}\right)$.

7.6 EXAMPLES Identifying critical points

7.6.1 $f(x) = x^3 - 2x^2 + x + 7$

To identify critical points we start by solving the derivative. Which is done with the power rule:

$$f(x) = x^3 - 2x^2 + x + 7 \tag{7.1}$$
$$f'(x) = 3x^2 - 4x + 1 \tag{7.2}$$

Next we need to find the point where f'(x) = 0. This is our critical point. To do this we will use the quadratic formula since we are working with a quadratic formula:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \tag{7.3}$$
$$a = 3, b = -4, c = 1 \tag{7.4}$$

We now insert the values into the quadratic formula:

$$x = \frac{-(-4) \pm \sqrt{(-4)^2 - 4 \cdot 3 \cdot 1}}{2 \cdot 3} \tag{7.5}$$
$$x = \frac{4 \pm \sqrt{16 - 12}}{6} \tag{7.6}$$
$$x = \frac{4 \pm \sqrt{4}}{6} \tag{7.7}$$
$$x = \frac{4 \pm 2}{6} \tag{7.8}$$
$$\tag{7.9}$$

We have now reduced the function, we can now solve the two solutions:

$$x = \frac{4+2}{6} = 1 \tag{7.10}$$
$$x = \frac{4-2}{6} = \frac{2}{6} = \frac{1}{3} \tag{7.11}$$
$$\tag{7.12}$$

These are our two critical points for the function $f(x) = x^3 - 2x^2 + x + 7$
Next we evaluate at each critical point by first computing the second derivative:

$$f'(x) = 3x^2 - 4x + 1 \tag{7.13}$$
$$f''(x) = 6x - 4 \tag{7.14}$$

We can now insert our critical points to evaluate them:

$$f''(1) = 6 \cdot 1 - 4 = 2 \tag{7.15}$$
$$f''(\frac{1}{3}) = 6 \cdot \frac{1}{3} - 4 = -2 \tag{7.16}$$

Based on this we can conclude that:

- the critical point 1 is a local minimum since f"(1) $>$ 0
- the critical point $\frac{1}{3}$ is a local maximium since $f''(\frac{1}{3}) < 0$

7.6.2 $f(x) = x^n \ln x, \quad n \in \mathbb{N} \setminus \{0\}$

To solve this we need to apply the product rule:

$$f'(x) = h'(x) \cdot g(x) + h(x) \cdot g'(x) \tag{7.17}$$
$$h(x) = x^n \tag{7.18}$$
$$h'(x) = (n-1)x^{n-1} \tag{7.19}$$
$$g(x) = ln(x) \tag{7.20}$$
$$g'(x) = \frac{1}{x} \tag{7.21}$$
$$\tag{7.22}$$

We now insert these values into the product rule:

$$f'(x) = n \cdot x^{n-1} \cdot ln(x) + x^n \cdot \frac{1}{x} \tag{7.23}$$
$$f'(x) = n \cdot x^{n-1} \cdot ln(x) + \frac{x^n}{x} \tag{7.24}$$
$$f'(x) = n \cdot x^{n-1} \cdot ln(x) + x^{n-1} \tag{7.25}$$

We can now reduce it by factoring the expression:

$$f'(x) = x^{n-1} \cdot (n \cdot ln(x) + 1) \tag{7.26}$$

We can now solve for f'(x) = 0:

$$\begin{align}
0 &= x^{n-1} \cdot (n \cdot ln(x) + 1) \tag{7.27}\\
\frac{0}{x^{n-1}} &= \frac{x^{n-1} \cdot (n \cdot ln(x) + 1)}{x^{n-1}}((n-1) \tag{7.28}\\
\frac{0}{n} &= \frac{n \cdot ln(x) + 1}{n} \tag{7.29}\\
0 &= ln(x) + \frac{1}{n} \tag{7.30}\\
ln(x) &= -\frac{1}{n} \tag{7.31}\\
e^{ln(x)} &= e^{-\frac{1}{n}} \tag{7.32}\\
x &= e^{-\frac{1}{n}} \tag{7.33}\\
& \tag{7.34}
\end{align}$$

We have now found the critical point for the function $f(x) = x^n \ln x, \quad n \in \mathbb{N} \setminus \{0\}$
Next we need to compute the second derivative of f'(x):

$$\begin{align}
f'(x) &= n \cdot x^{n-1} \cdot ln(x) + x^{n-1} \tag{7.35}\\
& \tag{7.36}
\end{align}$$

To compute this we again use the product rule and combine it with the sum rule:

$$\begin{align}
t(x) &= nx^{n-1} \tag{7.37}\\
t'(x) &= (n-1)nx^{n-2} \tag{7.38}\\
g(x) &= ln(x) \tag{7.39}\\
g'(x) &= \frac{1}{x} \tag{7.40}\\
h(x) &= x^{n-1} \tag{7.41}\\
h'(x) &= (n-1)x^{n-2} \tag{7.42}
\end{align}$$

We can now combine these to get the second derivative:

$$\begin{align}
f''(x) &= t'(x)g(x) + t(x)g'(x) + h'(x) \tag{7.43}\\
f''(x) &= (n-1)nx^{n-2} \cdot ln(x) + nx^{n-1} \cdot x^{n-1} + (n-1)x^{n-2} \tag{7.44}\\
f''(x) &= (n-1)nx^{n-2} \cdot ln(x) + nx^{n-2} + (n-1)x^{n-2} \tag{7.45}\\
& \tag{7.46}
\end{align}$$

Based on this we can conclude that since n>0 and $e^{-\frac{1}{n}} > 0$, then f"(x) > 0 and therefore a local minimum

7.6.3 $f(x, y) = 3x^2 + 2xy + 2y^2 - 6$

To solve this we will need to use partial differentiation, where we start by treating x as a variable and y as the constant. To do this we can simply use the power and sum rule:

$$\begin{align}
\frac{\partial f}{\partial x} &= 6x + 2y \tag{7.47}\\
\frac{\partial f}{\partial y} &= 2x + 4y \tag{7.48}\\
& \tag{7.49}
\end{align}$$

Now to solve this we first need to solve y using the first partial derivative:

$$0 = 6x + 2y \tag{7.50}$$
$$-\frac{2y}{-2} = \frac{6x}{-2} \tag{7.51}$$
$$y = -3x \tag{7.52}$$

We can now use that to solve $\frac{\partial f}{\partial y} = 0$ by inserting $y = -3x$:

$$0 = 2x + 4 \cdot -3x \tag{7.53}$$
$$0 = 2x - 12x \tag{7.54}$$
$$\frac{0}{-10} = -\frac{10x}{-10} \tag{7.55}$$
$$x = 0 \tag{7.56}$$
$$\tag{7.57}$$

We can now use x = 0 to solve $y = -3x$:

$$y = -3 \cdot 0 \tag{7.58}$$
$$y = 0 \tag{7.59}$$

This then gives us the following critical points:

$$(0, 0) \tag{7.60}$$

We can now evaluate the critical points. Since we now have two variables we need to use a Hessian Matrix and therefore need to compute the following points:

$$H(f)(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} \tag{7.61}$$

$$\frac{\partial^2 f}{\partial x^2} = 6 \tag{7.62}$$
$$\frac{\partial^2 f}{\partial y^2} = 4 \tag{7.63}$$
$$\frac{\partial^2 f}{\partial xy} = 2 \tag{7.64}$$
$$\frac{\partial^2 f}{\partial yx} = 2 \tag{7.65}$$
$$\tag{7.66}$$

We can now insert these values into the hessian matrix:

$$H(f)(x,y) = \begin{bmatrix} 6 & 2 \\ 2 & 4 \end{bmatrix} \tag{7.67}$$

We can now compute the Hessian Matrix:

$$D = 6 \cdot 4 - 2 \cdot 2 = 20 \tag{7.68}$$

Since D > 0, we know that the Hessian matrix is either a local minimum or a local maximum. We therefore check if the two double differentiated partial derivatives have the same sign:

$$\frac{\partial^2 f}{\partial x^2} = 6 \tag{7.69}$$
$$\frac{\partial^2 f}{\partial y^2} = 4 \tag{7.70}$$

They both share the same sign and they're both positive, we can therefore conclude that the critical point:

$$(0, 0) \tag{7.71}$$

is a local minimum

7.6.4 $f(x, y) = 3x^2 + 4xy + y^2 - 2x + 4y$

To solve this we use the same approach as before, starting by solving the partial derivatives:

$$\frac{\partial f}{\partial x} = 6x + 4y - 2 \tag{7.72}$$
$$\frac{\partial f}{\partial y} = 4x + 2y + 4 \tag{7.73}$$
$$\tag{7.74}$$

We now solve $\frac{\partial f}{\partial x} = 0$:

$$0 = 6x + 4y - 2 \tag{7.75}$$
$$2 = 6x + 4y \tag{7.76}$$
$$\frac{2}{2} = \frac{6x + 4y}{2} \tag{7.77}$$
$$1 = 3x + 2y \tag{7.78}$$
$$\tag{7.79}$$

Next we do the same for $\frac{\partial f}{\partial y} = 0$

$$0 = 4x + 2y + 4 \tag{7.80}$$
$$-4 = 4x + 2y \tag{7.81}$$
$$-\frac{4}{2} = \frac{4x + 2y}{2} \tag{7.82}$$
$$-2 = 2x + y \tag{7.83}$$
$$\tag{7.84}$$

We can now start finding the critical points. We will start by solving y, using $-2 = 2x + y$:

$$y = -2 - 2x \tag{7.85}$$

We can now use $y = -2 - 2x$ to find x by inserting it into the function $1 = 3x + 2y$:

$$1 = 3x + 2(-2 - 2x) \tag{7.86}$$
$$1 = 3x + -4 - 4x \tag{7.87}$$
$$1 + 4 = -x \tag{7.88}$$
$$5 = -x \tag{7.89}$$
$$x = -5 \tag{7.90}$$
$$\tag{7.91}$$

We can now use $x = -5$ to solve y:

$$
\begin{aligned}
y &= -2 - 2 \cdot -5 && (7.92)\\
y &= -2 + 10 && (7.93)\\
y &= 8 && (7.94)\\
& && (7.95)
\end{aligned}
$$

Therefore our critical points are at

$$(-5, 8) \tag{7.96}$$

We know compute the Hessian matrix:

$$
\begin{aligned}
\frac{\partial^2 f}{\partial x^2} &= 6 && (7.97)\\
\frac{\partial^2 f}{\partial y^2} &= 2 && (7.98)\\
\frac{\partial^2 f}{\partial xy} &= 4 && (7.99)\\
\frac{\partial^2 f}{\partial yx} &= 4 && (7.100)\\
& && (7.101)
\end{aligned}
$$

We can now insert these values into the hessian matrix:

$$H(f)(x, y) = \begin{bmatrix} 6 & 4 \\ 4 & 2 \end{bmatrix} \tag{7.102}$$

We can now compute the Hessian Matrix:

$$D = 6 \cdot 2 - 4 \cdot 4 = -4 \tag{7.103}$$

Since the product of the Hessian matrix < 0, we know that is a straddle point at the critical point $(-5, 8)$

8 Integrals

8.1 Common integrals

f	$\int f\,dx$	f	$\int f\,dx$
x^n	$\frac{x^{n+1}}{n+1}+C$	$\ln x$	$x\ln x - x + C$
e^x	$e^x + C$	a^x	$\frac{a^x}{\ln a}+C$
e^{ax}	$\frac{e^{ax}}{a}+C$	$\sin x$	$-\cos x + C$
$\cos x$	$\sin x + C$	$-\sin x$	$\cos x + C$
$-\cos x$	$-\sin x + C$	$\tan x$	$-\ln\lvert\cos x\rvert + C$
$\sqrt{x}$	$\frac{2}{3}x^{3/2}+C$	$\arctan x$	$x\arctan x - \frac{1}{2}\ln\left(1+x^2\right)+C$
$\arcsin x$	$x\arcsin x + \sqrt{1-x^2}+C$	$\arccos x$	$x\arccos x - \sqrt{1-x^2}+C$
f^n	depends on f	$\frac{1}{f}$	depends on f
e^f	$e^f\cdot f' + C$	$\ln f$	$\int \ln f\,dx + C$
$\sinh x$	$\cosh x + C$	$\cosh x$	$\sinh x + C$
$\tanh x$	$\ln\lvert\cosh x\rvert + C$	$\coth x$	$\ln\lvert\sinh x\rvert + C$
$\sec x$	$\ln\lvert\sec x + \tan x\rvert + C$	$\log_b x$	$\frac{x}{\ln b}(\log_b x - 1) + C$

8.2 Common derivatives

Useful for u-substitution and partial integration,

f	f'	f	f'
x^n	nx^{n-1}	$\ln x$	$\frac{1}{x}$
e^x	e^x	a^x	$\ln a\cdot a^x$
e^{ax}	ae^{ax}	$\sin x$	$\cos x$
$\cos x$	$-\sin x$	$-\sin x$	$-\cos x$
$-\cos x$	$\sin x$	$\tan x$	$\sec^2 x$
$\sqrt{x}$	$\frac{1}{2\sqrt{x}}$	$\arctan x$	$\frac{1}{1+x^2}$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$	$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$
f^n	$nf'f^{n-1}$	$\frac{1}{f}$	$-\frac{f'}{f^2}$
e^f	$f'e^f$	$\ln f$	$\frac{f'}{f}$
$\sinh x$	$\cosh x$	$\cosh x$	$\sinh x$
$\tanh x$	$\operatorname{sech}^2 x$	$\coth x$	$-\operatorname{csch}^2 x$
$\sec x$	$\sec x\tan x$	$\log_b x$	$\frac{1}{x\ln b}$

8.3 Constant Rule for Integration

Integrating a constant k over x simply multiplies k by x, reflecting the area under a horizontal line.

$$\int k\,dx = kx + C \quad \text{(where } k \text{ is a constant)}$$

Examples:

- $\int 5\,dx = 5x + C$
- $\int -4\,dx = -4x + C$
- $\int 0\,dx = C$ (represents the integral of zero, which is a constant)

8.4 Power Rule for Integration

The power rule for integration is applicable except when the exponent is -1, which instead results in a logarithmic function.

$$\int x^n\,dx = \frac{1}{n+1}x^{n+1} + C \quad \text{(where } n \neq -1\text{)}$$

Examples:

- $\int x^2\,dx = \frac{1}{3}x^3 + C$

- $\int x^4 \, dx = \frac{1}{5}x^5 + C$
- $\int x^{-2} \, dx = -x^{-1} + C = -\frac{1}{x} + C$ (integration of x to the negative two)

8.5 Exponential Function Integration

The exponential function e^x maintains its form upon integration, which reflects its unique property of being equal to its own derivative.

$$\int e^x \, dx = e^x + C$$

Examples:

- $\int 2e^x \, dx = 2e^x + C$ (integration of two times the exponential function)
- $\int e^{2x} \, dx = \frac{1}{2}e^{2x} + C$ (requires a simple adjustment for the coefficient inside the exponent)
- $\int 3e^{-x} \, dx = -3e^{-x} + C$ (integrating three times the exponential decay function)

8.6 Integration of $\frac{1}{x}$

The integral of $\frac{1}{x}$ results in the natural logarithm of the absolute value of x, highlighting the logarithmic integration rule.

$$\int \frac{1}{x} \, dx = \ln|x| + C$$

Examples:

- $\int \frac{2}{x} \, dx = 2\ln|x| + C$ (scaling the logarithmic output by two)
- $\int \frac{1}{2x} \, dx = \frac{1}{2}\ln|x| + C$ (the coefficient affects the logarithm's multiplier)
- $\int \frac{1}{3x} \, dx = \frac{1}{3}\ln|x| + C$ (similar effect with a different scaling factor)

8.7 Trigonometric Functions Integration

8.7.1 Cosine Function Integration

The integral of the cosine function, due to its periodic nature, results in the sine function.

$$\int \cos(x) \, dx = \sin(x) + C$$

Examples:

- $\int 2\cos(x) \, dx = 2\sin(x) + C$ (Scaling the integral by a factor of 2.)
- $\int \cos(2x) \, dx = \frac{1}{2}\sin(2x) + C$ (Applying the chain rule, factor of $\frac{1}{2}$ comes from the derivative of the inner function $2x$.)
- $\int 3\cos(3x) \, dx = \sin(3x) + C$ (Here, $\frac{1}{3}$ from the derivative of $3x$ is offset by the factor of 3 outside.)

8.7.2 Sine Function Integration

Similarly, the integral of the sine function results in the negative cosine function, reflecting the phase shift in trigonometric identities.

$$\int \sin(x) \, dx = -\cos(x) + C$$

Examples:

- $\int 2\sin(x) \, dx = -2\cos(x) + C$ (Scaling affects the amplitude of the cosine function.)
- $\int \sin(2x) \, dx = -\frac{1}{2}\cos(2x) + C$ (The $\frac{1}{2}$ compensates for the faster oscillation of $\cos(2x)$.)
- $\int 3\sin(3x) \, dx = -\cos(3x) + C$ (The factor of $\frac{1}{3}$ from the derivative of $3x$ is multiplied by 3.)

8.8 Logarithmic Function Integration

The integration of logarithmic functions typically involves the product of x with the natural logarithm minus the integral of x.

$$\int \frac{1}{x}\,dx = \ln|x| + C$$

$$\int \ln(x)\,dx = x\ln(x) - x + C$$

Examples:

- $\int 2\ln(x)\,dx = 2x\ln(x) - 2x + C$ (Doubling the integral of $\ln(x)$.)
- $\int \ln(2x)\,dx = x\ln(2x) - x + C$ (Includes using $\ln(ab) = \ln(a) + \ln(b)$.)
- $\int 3\ln(3x)\,dx = 3x\ln(3x) - 3x + C$ (Scaling by 3 and using properties of logarithms.)

8.9 Square Root Function Integration

The integration of functions involving square roots can often be simplified by recalling the power rule in its general form $\int x^n\,dx = \frac{x^{n+1}}{n+1} + C$.

$$\int \sqrt{x}\,dx = \int x^{1/2}\,dx = \frac{2}{3}x^{3/2} + C$$

$$\int \frac{1}{\sqrt{x}}\,dx = \int x^{-1/2}\,dx = 2\sqrt{x} + C$$

Examples:

- $\int 2\sqrt{x}\,dx = 2\cdot\frac{2}{3}x^{3/2} = \frac{4}{3}x^{3/2} + C$
- $\int 3\sqrt{x}\,dx = 3\cdot\frac{2}{3}x^{3/2} = 2x^{3/2} + C$
- $\int 4\frac{1}{\sqrt{x}}\,dx = 4\cdot 2\sqrt{x} = 8\sqrt{x} + C$

8.10 Sum and Difference Rule for Integration

The sum and difference rule is a direct application of the linearity of integrals, allowing the integration of each term separately.

$$\int (f(x) \pm g(x))\,dx = \int f(x)\,dx \pm \int g(x)\,dx$$

Examples:

- $\int (x^2 + x)\,dx = \int x^2\,dx + \int x\,dx = \frac{1}{3}x^3 + \frac{1}{2}x^2 + C$
- $\int (3x^3 - 4x)\,dx = \int 3x^3\,dx - \int 4x\,dx = \frac{3}{4}x^4 - 2x^2 + C$
- $\int (5x^2 + 6x - 7)\,dx = \int 5x^2\,dx + \int 6x\,dx - \int 7\,dx = \frac{5}{3}x^3 + 3x^2 - 7x + C$
- $\int (x^4 - 2x^2 + x)\,dx = \int x^4\,dx - \int 2x^2\,dx + \int x\,dx = \frac{1}{5}x^5 - \frac{2}{3}x^3 + \frac{1}{2}x^2 + C$

8.11 Integration by Substitution (u-Substitution)

Integration by substitution, commonly known as u-substitution, is a method used to simplify integrals by changing the variable of integration. This technique is especially useful when integrating composite functions or functions with an apparent inner function and its derivative.
When to use: Use u-substitution when the integral contains a function and its derivative, or when a substitution can simplify the integral into a familiar form.
Example: Integrate $\int x\cos(x^2)\,dx$.
Step 1: Identify the inner function and its differential.

$$u = x^2, \qquad du = 2x\,dx$$

Step 2: Solve for dx.

$$2x\,dx = du \quad \Rightarrow \quad x\,dx = \frac{1}{2}du$$

Step 3: Substitute and integrate.

$$\begin{aligned}\int x\cos(x^2)\,dx &= \int \cos(u)\cdot\frac{1}{2}du = \frac{1}{2}\int\cos(u)\,du \\ &= \frac{1}{2}\sin(u) + C \\ &= \frac{1}{2}\sin(x^2) + C\end{aligned}$$

8.11.1 Additional Examples:

Example: Integrate $\int e^{3x}\,dx$.
Step 1: Identify the inner function and its differential.

$$u = 3x, \qquad du = 3\,dx$$

Step 2: Solve for dx.

$$dx = \frac{du}{3}$$

Step 3: Substitute and integrate.

$$\int e^u \cdot \frac{du}{3} = \frac{1}{3}\int e^u\,du = \frac{1}{3}e^u + C = \frac{1}{3}e^{3x} + C$$

Example: Integrate $\int \sin(\sqrt{x})\,dx$.
Step 1: Identify the inner function and its differential.

$$u = \sqrt{x}, \qquad du = \frac{1}{2\sqrt{x}}\,dx$$

Step 2: Solve for dx.

$$dx = 2u\,du$$

Step 3: Substitute and integrate.

$$\int \sin(u)\cdot 2u\,du = 2\int u\sin(u)\,du$$

This integral can be further solved by integration by parts.

Example: Integrate $\int \frac{1}{x\ln(x)}\,dx$.
Step 1: Identify the inner function and its differential.

$$u = \ln(x), \qquad du = \frac{1}{x}\,dx$$

Step 2: Substitute and integrate.

$$\int \frac{1}{x\ln(x)}\,dx = \int \frac{1}{u}\,du = \ln|u| + C = \ln|\ln(x)| + C$$

Example: Integrate $\int \tan(x)\sec^2(x)\,dx$.
Step 1: Identify the inner function and its differential.

$$u = \tan(x), \qquad du = \sec^2(x)\,dx$$

Step 2: Substitute and integrate.

$$\int u\,du = \frac{u^2}{2} + C = \frac{\tan^2(x)}{2} + C$$

8.12 Integration by Parts

The formula for integration by parts is given by:

$$\int u\,dv = uv - \int v\,du$$

This technique is typically used when an integral involves the product of two functions.

Example: Integrate $\int xe^x\,dx$.
Step 1: Identify u and dv.

$$u = x, \qquad dv = e^x\,dx$$

Step 2: Differentiate u and integrate dv.

$$du = dx, \qquad v = e^x$$

Step 3: Apply the integration by parts formula.

$$\int xe^x\,dx = xe^x - \int e^x\,dx = xe^x - e^x + C$$

Example: Integrate $\int x\ln(x)\,dx$.
Step 1: Identify u and dv.

$$u = \ln(x), \qquad dv = x\,dx$$

Step 2: Differentiate u and integrate dv.

$$du = \frac{1}{x}\,dx, \qquad v = \frac{x^2}{2}$$

Step 3: Apply the integration by parts formula.

$$\int x\ln(x)\,dx = \frac{x^2}{2}\ln(x) - \int \frac{x^2}{2}\cdot\frac{1}{x}\,dx = \frac{x^2}{2}\ln(x) - \frac{x^2}{4} + C$$

Additional Examples:
Example: Integrate $\int \sin(x)x^2\,dx$.
Step 1: Identify u and dv.

$$u = x^2, \qquad dv = \sin(x)\,dx$$

Step 2: Differentiate u and integrate dv.

$$du = 2x\,dx, \qquad v = -\cos(x)$$

Step 3: Apply the integration by parts formula.

$$\int x^2\sin(x)\,dx = -x^2\cos(x) + \int 2x\cos(x)\,dx$$

Example: Integrate $\int \ln(x)e^x\,dx$.
Step 1: Identify u and dv.

$$u = \ln(x), \qquad dv = e^x\,dx$$

Step 2: Differentiate u and integrate dv.

$$du = \frac{1}{x}\,dx, \qquad v = e^x$$

Step 3: Apply the integration by parts formula.

$$\int \ln(x)e^x\,dx = \ln(x)e^x - \int e^x\cdot\frac{1}{x}\,dx$$

8.13 Double Integrals

Examples:
Example 1: Compute the double integral of $f(x, y) = x + y$ over the rectangle defined by $0 \leq x \leq 2$ and $0 \leq y \leq 1$.

$$\int_0^1 \int_0^2 (x + y)\, dx\, dy$$

First, integrate with respect to x:

$$\int_0^2 (x + y)\, dx = \left[\frac{x^2}{2} + yx\right]_0^2 = 2 + 2y$$

Then, integrate with respect to y:

$$\int_0^1 (2 + 2y)\, dy = \left[2y + y^2\right]_0^1 = 3$$

Example 2: Evaluate the double integral of $f(x, y) = xy$ over the region bounded by $y = x^2$ and $y = 1$.

$$\int_0^1 \int_{x^2}^1 xy\, dy\, dx$$

First, integrate with respect to y:

$$\int_{x^2}^1 xy\, dy = x\left[\frac{y^2}{2}\right]_{x^2}^1 = x\left(\frac{1}{2} - \frac{x^4}{2}\right)$$

Then, integrate with respect to x:

$$\int_0^1 \left(\frac{x}{2} - \frac{x^5}{2}\right) dx = \left[\frac{x^2}{4} - \frac{x^6}{12}\right]_0^1 = \frac{1}{4} - \frac{1}{12} = \frac{1}{6}$$

8.13.1 Example with double integral and R

On the rectangle $R = \{(x, y) \mid a \leq x \leq b, c \leq y \leq d\}$, solve the integral

$$\iint_R (3xy^2 - xy + 1)\, dA.$$

Here dA means $dxdy = dydx$. Write it explicitly first with bounds on x and on y. Compute the volume between the xy-plane and $3xy^2 - xy + 1$ on the domain $R = [-1, 2] \times [3, 4]$. Include and explain the intermediary steps.
Solution
To solve the double integral over the specified rectangle, we first define the limits for each variable. The outer integral will be with respect to x, and the inner integral will be with respect to y. This setup allows us to integrate the function over the rectangular domain $R = [-1, 2] \times [3, 4]$, where x ranges from -1 to 2, and y ranges from 3 to 4:
Setup the Integral
Define the limits for x and y:

$$\int_{-1}^2 \int_3^4 (3xy^2 - xy + 1)\, dy\, dx$$

Integrate with Respect to y
First, we integrate the inner integral with respect to y:

$$\begin{aligned}
\int_3^4 (3xy^2 - xy + 1)\, dy &= 3x \int_3^4 y^2\, dy - x \int_3^4 y\, dy + \int_3^4 1\, dy \\
&= 3x\left[\frac{y^3}{3}\right]_3^4 - x\left[\frac{y^2}{2}\right]_3^4 + [y]_3^4 \\
&= x\left(\frac{64}{3} - \frac{27}{3}\right) - x\left(\frac{16}{2} - \frac{9}{2}\right) + (4 - 3) \\
&= x\left(\frac{37}{3}\right) - x\left(\frac{7}{2}\right) + 1 \\
&= \frac{37x}{3} - \frac{7x}{2} + 1.
\end{aligned}$$

Integrate with Respect to x
Next, integrate the resulting expression with respect to x:

$$\begin{aligned}\int_{-1}^{2}\left(\frac{37x}{3}-\frac{7x}{2}+1\right)dx &= \frac{37}{3}\left[\frac{x^2}{2}\right]_{-1}^{2}-\frac{7}{2}\left[\frac{x^2}{2}\right]_{-1}^{2}+[x]_{-1}^{2} \\ &= \frac{37}{3}\left(\frac{4}{2}-\frac{1}{2}\right)-\frac{7}{2}\left(\frac{4}{2}-\frac{1}{2}\right)+(2+1) \\ &= \frac{37}{3}\cdot\frac{3}{2}-\frac{7}{2}\cdot\frac{3}{2}+3 \\ &= \frac{111}{2}-\frac{21}{2}+3 \\ &= \frac{90}{2}+3 \\ &= 45+3 \\ &= 48\end{aligned}$$

Therefore, the result of the double integral is:

$$\int_{-1}^{2}\int_{3}^{4}(3xy^2-xy+1)\,dy\,dx = 48$$

8.14 Trigonometric Integrals:

$\int \sec^2(x)\,dx = \tan(x)+C$. Example: $\int 3\sec^2(x)\,dx = 3\tan(x)+C$
$\int \csc^2(x)\,dx = -\cot(x)+C$. Example: $\int 2\csc^2(x)\,dx = -2\cot(x)+C$
$\int \sec(x)\tan(x)\,dx = \sec(x)+C$. Example: $\int 2\sec(x)\tan(x)\,dx = 2\sec(x)+C$
$\int \csc(x)\cot(x)\,dx = -\csc(x)+C$. Example: $\int 3\csc(x)\cot(x)\,dx = -3\csc(x)+C$

8.15 Exponential and Logarithmic Integrals

Exponential integrals typically involve functions of the form a^x where a is a positive constant different from 1. The base a influences the rate of growth or decay of the function.

$$\int a^x\,dx = \frac{a^x}{\ln(a)}+C$$

Example: $\int 2^x\,dx = \frac{2^x}{\ln(2)}+C$
Additional Examples:

- $\int 10^x\,dx = \frac{10^x}{\ln(10)}+C$
- $\int e^{kx}\,dx = \frac{1}{k}e^{kx}+C$ (where k is a non-zero constant)

Example for e^{kx}: $\int e^{3x}\,dx = \frac{1}{3}e^{3x}+C$

8.16 Inverse Trigonometric Integrals

Integrals involving inverse trigonometric functions are commonly encountered in problems dealing with triangles and angular relationships.

$$\int \frac{1}{\sqrt{1-x^2}}\,dx = \arcsin(x)+C$$

Example: $\int \frac{1}{\sqrt{1-x^2}}\,dx = \arcsin(x)+C$
Additional Example:

$$\int \frac{1}{\sqrt{1-x^2}}\,dx = \arccos(x)+C \quad \left(\text{since } \arccos(x)=\frac{\pi}{2}-\arcsin(x)\right)$$

$$\int \frac{1}{1+x^2}\,dx = \arctan(x)+C$$

Example: $\int \frac{1}{1+x^2}\,dx = \arctan(x)+C$

8.17 Partial Fractions

Partial fractions decomposition is a technique used to simplify the integration of rational functions where the degree of the numerator is less than the degree of the denominator. By decomposing a complex fraction into simpler fractions, integration can often be reduced to a sum of simpler, more standard integrals.
Example: Consider the integral:

$$\int \frac{2x+3}{(x+1)(x-2)}\,dx$$

To integrate this, first decompose the fraction:

$$\frac{2x+3}{(x+1)(x-2)} = \frac{A}{x+1} + \frac{B}{x-2}$$

Solve for A and B by equating coefficients:

$$2x+3 = A(x-2) + B(x+1)$$

Setting $x=2$ yields $B=1$, and setting $x=-1$ yields $A=1$. Therefore, the integral simplifies to:

$$\int \left(\frac{1}{x+1} + \frac{1}{x-2}\right) dx = \ln|x+1| - \ln|x-2| + C$$

Additional Examples:
Example: Decompose and integrate $\int \frac{x^2+2x+3}{(x-1)(x^2+1)}\,dx$. Decompose into partial fractions:

$$\frac{x^2+2x+3}{(x-1)(x^2+1)} = \frac{A}{x-1} + \frac{Bx+C}{x^2+1}$$

Matching coefficients, solve for A, B, and C to find $A=1$, $B=1$, and $C=1$. Then:

$$\int \left(\frac{1}{x-1} + \frac{x+1}{x^2+1}\right) dx = \ln|x-1| + \frac{1}{2}\ln\left(x^2+1\right) + \arctan(x) + C$$

Example: Decompose and integrate $\int \frac{3x+4}{x^2-4}\,dx$. The fraction can be rewritten as:

$$\frac{3x+4}{x^2-4} = \frac{3x+4}{(x-2)(x+2)} = \frac{A}{x-2} + \frac{B}{x+2}$$

Solving for A and B gives $A = \frac{3}{2}$ and $B = \frac{5}{2}$. The integral then becomes:

$$\int \left(\frac{\frac{3}{2}}{x-2} + \frac{\frac{5}{2}}{x+2}\right) dx = \frac{3}{2}\ln|x-2| + \frac{5}{2}\ln|x+2| + C$$

8.18 Definite Integrals

Definite integrals compute the net area under a curve between two points and are evaluated by finding the antiderivative and applying the Fundamental Theorem of Calculus.
Example:

$$\int_0^1 x^2\,dx = \left[\frac{1}{3}x^3\right]_0^1 = \frac{1}{3} - 0 = \frac{1}{3}$$

Additional Examples:

$$\int_{-1}^1 x^3\,dx = \left[\frac{1}{4}x^4\right]_{-1}^1 = \frac{1}{4} - \frac{1}{4} = 0 \quad \text{(odd function over a symmetric interval)}$$

$$\int_0^\pi \sin(x)\,dx = [-\cos(x)]_0^\pi = -\cos(\pi) - (-\cos(0)) = 2$$

Example: Integrate $\int_{-1}^1 x^3\,dx$.

$$\int_{-1}^1 x^3\,dx = \left[\frac{1}{4}x^4\right]_{-1}^1 = \frac{1}{4}(1^4) - \frac{1}{4}((-1)^4) = 0 \quad \text{(odd function over a symmetric interval)}$$

Example: Integrate $\int_0^\pi \sin(x)\,dx$.

$$\int_0^\pi \sin(x)\,dx = [-\cos(x)]_0^\pi = -\cos(\pi) + \cos(0) = 2$$

Example: Compute the integral of $\int_0^2 3x^2\,dx$.

$$\int_0^2 3x^2\,dx = \left[x^3\right]_0^2 = 2^3 - 0^3 = 8$$

Example: Evaluate the area under the curve of $\int_1^4 \frac{1}{x}\,dx$.

$$\int_1^4 \frac{1}{x}\,dx = [\ln|x|]_1^4 = \ln(4) - \ln(1) = \ln(4)$$

Example: Determine the total area between $\int_0^{2\pi} \cos(x)\,dx$.

$$\int_0^{2\pi} \cos(x)\,dx = [\sin(x)]_0^{2\pi} = \sin(2\pi) - \sin(0) = 0$$

Example: Calculate the integral of $\int_{-2}^2 \sqrt{4-x^2}\,dx$, which represents the area of a semicircle with radius 2.

$$\int_{-2}^2 \sqrt{4-x^2}\,dx = \pi \times 2^2 \times \frac{1}{2} = 2\pi$$

8.19 Leibniz Rule for Integration

The Leibniz rule allows us to differentiate an integral with respect to a parameter, particularly useful when the limits of the integral depend on the variable of differentiation. If we have an integral of the form

$$F(x) = \int_{g(x)}^{h(x)} f(u)\,du,$$

then the derivative with respect to x is given by:

$$F'(x) = f(h(x)) \cdot h'(x) - f(g(x)) \cdot g'(x).$$

This rule states that we evaluate the integrand at the upper and lower limits, multiply by the derivative of each limit, and subtract the lower term from the upper term.
Examples:

- **Example 1:** Differentiate $f(x) = \int_x^\infty \left(\sqrt{3u} + \cos^2(u) + 3\right)\,du$.

  $$f'(x) = -\left(\sqrt{3x} + \cos^2(x) + 3\right)$$

  Here, the upper limit is ∞, which has a derivative of zero, so only the lower limit contributes to the result.

- **Example 2:** Differentiate $f(x) = \int_x^a \left(\sqrt{3u} + \cos^2(u) + 3\right)\,du$ where a is a constant.

  $$f'(x) = -\left(\sqrt{3x} + \cos^2(x) + 3\right)$$

  Since the upper limit a is constant, only the lower limit x affects the derivative.

- **Example 3:** Differentiate $f(x) = \int_{3x}^{2x} \frac{u^2-1}{u^2+1}\,du$.

  $$f'(x) = \frac{2(4x^2-1)}{4x^2+1} - \frac{3(9x^2-1)}{9x^2+1}$$

  Here, both limits depend on x, so we apply the rule to both terms. The derivative at the upper limit is multiplied by $h'(x) = 2$, and the derivative at the lower limit is multiplied by $g'(x) = 3$.

- **Example 4:** Differentiate $F(x) = \int_x^{x^2} e^{u^2}\, du$.

$$F'(x) = e^{(x^2)^2} \cdot 2x - e^{x^2} \cdot 1 = 2x\, e^{x^4} - e^{x^2}$$

  The upper limit is $h(x) = x^2$ with $h'(x) = 2x$, and the lower limit is $g(x) = x$ with $g'(x) = 1$. Applying the rule gives us the derivative.

- **Example 5:** Differentiate $G(x) = \int_0^{\sin(x)} \frac{1}{1+u^2}\, du$.

$$G'(x) = \frac{1}{1 + (\sin(x))^2} \cdot \cos(x)$$

  Here, only the upper limit $h(x) = \sin(x)$ depends on x, so we evaluate the integrand at $u = \sin(x)$ and multiply by $h'(x) = \cos(x)$.

8.20 Primitives (Antiderivatives) of Hyperbolic Functions

The primitive, or antiderivative, of a function $f(x)$ is a function $F(x)$ such that $F'(x) = f(x)$. Finding primitives is essential in integration, where we reverse the differentiation process.
For hyperbolic functions, the following primitives hold:

$$\int \cosh(x)\, dx = \sinh(x) + C$$

$$\int \sinh(x)\, dx = \cosh(x) + C$$

$$\int \tanh(x)\, dx = \ln|\cosh(x)| + C$$

8.20.1 Examples

- $\int \cosh(x)\, dx = \sinh(x) + C$
  The hyperbolic cosine function integrates to the hyperbolic sine function.
- $\int \sinh(x)\, dx = \cosh(x) + C$
  The hyperbolic sine function integrates to the hyperbolic cosine function.
- $\int \tanh(x)\, dx = \ln|\cosh(x)| + C$
  The hyperbolic tangent function integrates to the natural logarithm of $\cosh(x)$.

These antiderivatives are useful when dealing with hyperbolic functions in calculus, providing a basis for solving integrals involving $\cosh(x)$, $\sinh(x)$, and $\tanh(x)$.

ordered sampling with replacement	n^k
ordered sampling without replacement	$P_k^n = \frac{n!}{(n-k)!}$
unordered sampling without replacement	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$
unordered sampling with replacement	$\binom{n+k-1}{k}$

Figure 4: Counting methods

9 Sampling

Sampling methods are used to select a subset of elements from a larger population. The chosen method depends on whether the order of elements matters and whether elements can be chosen more than once. We discuss four primary methods: ordered and unordered sampling, both with and without replacement.

9.1 Conditional Probability

Conditional probability measures the probability of an event occurring given that another event has already occurred. It is defined as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

where $P(A|B)$ is the probability of event A given event B, $P(A \cap B)$ is the probability of both A and B occurring, and $P(B)$ is the probability of B occurring.
Example: Suppose there is a bag of 10 marbles, 3 of which are red. If a red marble is drawn and not replaced, what is the probability of drawing another red marble?

$$P(R_2|R_1) = \frac{P(R_1 \cap R_2)}{P(R_1)} = \frac{\frac{3}{10} \times \frac{2}{9}}{\frac{3}{10}} = \frac{2}{9}$$

9.2 Independence

Two events A and B are independent if the occurrence of A does not affect the probability of B occurring, and vice versa. Mathematically, two events are independent if:

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

or equivalently,

$$P(A|B) = P(A) \quad \text{and} \quad P(B|A) = P(B)$$

Example: Consider flipping a fair coin and rolling a fair six-sided die. The outcome of the coin toss does not affect the outcome of the die roll. Therefore, these two events are independent. If A is the event "heads" and B is the event "rolling a 4," then:

$$P(A) = \frac{1}{2}, \quad P(B) = \frac{1}{6}, \quad P(A \cap B) = P(A)P(B) = \frac{1}{2} \times \frac{1}{6} = \frac{1}{12}$$

9.3 Selecting n and k

9.3.1 Selecting n

- **Population Size:** n should represent the total number of distinct elements or outcomes available for selection. For instance, if you are drawing cards from a deck, n would be 52.
- **Scope of Study:** In studies where a subset of the population is to be analyzed, n should reflect only the elements relevant to the study.

9.3.2 Selecting k

- **Objective of Sampling:** k should align with the objective of your analysis. For example, if the goal is to form teams or committees, k would be the number of positions or slots to fill. So if 5 positions needs to be filled, then k = 5

9.3.3 Practical Examples

- In a marketing survey, if a company can only afford to survey 20 people out of a customer base of 100, $n = 100$ and $k = 20$.
- In a lottery draw, if there are 50 numbers and 6 are to be chosen, $n = 50$ and $k = 6$.

9.4 Ordered Sampling with Replacement

In this method, each element can be chosen multiple times with the order of selection being significant.

$$n^k$$

Examples:

1. Choosing 2 digits to create a two-digit number from the digits 0 through 9. Here, $n = 10$ and $k = 2$:

$$10^2 = 100 \text{ possible numbers}$$

2. Throwing 3 dice together. Here n = 6 (6 sides to dice) and k = 3

$$6^3 = 216\text{possible throws}$$

3. Selecting moves in a game where each move can be one of 5 options, repeated across 3 moves. Here, $n = 5$ and $k = 3$:

$$5^3 = 125 \text{ possible sequences of moves}$$

9.5 Ordered Sampling without Replacement

Each element is unique in the selection and the order of selection is important.

$$P_k^n = \frac{n!}{(n-k)!}$$

Examples:

1. Drawing 3 different cards from a deck of 52 without replacing any. Here, $n = 52$ and $k = 3$:

$$P_3^{52} = \frac{52!}{(52-3)!} = 132,600 \text{ ways}$$

2. Choosing 5 different books from a shelf of 10, without replacement. Here, $n = 10$ and $k = 5$:

$$P_5^{10} = \frac{10!}{(10-5)!} = 30,240 \text{ ways}$$

3. Selecting 4 different employees to present in sequential order from a team of 20. Here, $n = 20$ and $k = 4$:

$$P_4^{20} = \frac{20!}{(20-4)!} = 116,280 \text{ ways}$$

9.6 Unordered Sampling without Replacement

Selection does not consider order, and elements are not repeated.

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

Examples:

1. Forming a committee of 4 from 12 candidates. Here, $n = 12$ and $k = 4$:

$$\binom{12}{4} = \frac{12!}{4!8!} = 495 \text{ committees}$$

2. Picking 6 fruits from a basket of 10 different fruits. Here, $n = 10$ and $k = 6$:

$$\binom{10}{6} = \frac{10!}{6!4!} = 210 \text{ combinations}$$

3. Selecting 2 winners from a group of 8 contestants. Here, $n = 8$ and $k = 2$:

$$\binom{8}{2} = \frac{8!}{2!6!} = 28 \text{ pairs}$$

9.7 Unordered Sampling with Replacement

Elements can be chosen more than once, but the order of selection does not matter.

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

Examples:

1. Choosing 4 desserts from a menu of 6 options, where selections can repeat. Here, $n = 6$ and $k = 4$:

$$\binom{6+4-1}{4} = \binom{9}{4} = 126 \text{ ways}$$

2. Filling 3 identical vases with flowers from 5 varieties. Here, $n = 5$ and $k = 3$:

$$\binom{5+3-1}{3} = \binom{7}{3} = 35 \text{ combinations}$$

3. Selecting 3 ingredients for a recipe from 4 available spices, allowing repetition. Here, $n = 4$ and $k = 3$:

$$\binom{4+3-1}{3} = \binom{6}{3} = 20 \text{ mixtures}$$

Bayes' Theorem is a fundamental concept in probability theory that describes how to update the probability of a hypothesis based on new evidence. It is a powerful tool for statistical inference.

9.8 Bayes' Theorem

Bayes' Theorem relates conditional and marginal probabilities of events. It is expressed as:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

where:

- $P(A|B)$ is the probability of event A given that B has occurred.
- $P(B|A)$ is the probability of event B given that A has occurred.
- $P(A)$ and $P(B)$ are the probabilities of observing A and B independently.

9.9 Example 1: School Transportation

Consider the situation where a student, Ashish, can either take the bus or be driven to school by his father. The probabilities involved are as follows:

- $P(\text{Bus}) = 0.8$ (Probability that Ashish takes the bus)
- $P(\text{Father}) = 0.2$ (Probability that Ashish's father drives him)
- $P(\text{Late}|\text{Bus}) = 0.5$ (Probability of being late given that Ashish took the bus)
- $P(\text{Late}|\text{Father}) = 0.2$ (Probability of being late given that Ashish's father drove him)

If Ashish is late to school, we can use Bayes' Theorem to calculate the probability that he was driven by his father:

$$P(\text{Father}|\text{Late}) = \frac{P(\text{Late}|\text{Father})P(\text{Father})}{P(\text{Late})}$$

Where $P(\text{Late})$ is the total probability of Ashish being late, which includes both modes of transportation:

$$P(\text{Late}) = P(\text{Late}|\text{Bus})P(\text{Bus}) + P(\text{Late}|\text{Father})P(\text{Father}) = 0.5 \times 0.8 + 0.2 \times 0.2 = 0.44$$

Plugging in the values, we get:

$$P(\text{Father}|\text{Late}) = \frac{0.2 \times 0.2}{0.44} \approx 0.09$$

This indicates that there is approximately a 9% chance that Ashish was driven by his father on a day he was late to school.

9.9.1 Example 2: Medical Testing

Consider a rare disease where:

- The probability of having the disease ($P(D)$) is $\frac{1}{100}$.
- The probability of testing positive if having the disease ($P(T|D)$) is $\frac{9}{10}$.
- The probability of testing positive ($P(T)$), calculated by considering both false positives and true positives, is $\frac{1}{10}$.

Calculate the probability of having the disease given a positive test result ($P(D|T)$):

$$P(D|T) = \frac{P(T|D)P(D)}{P(T)} = \frac{\frac{9}{10} \times \frac{1}{100}}{\frac{1}{10}} = \frac{9}{100} = 0.09 \text{ or } 9\%$$

9.10 Example 3: Spam Email Filtering

Suppose:

- The probability of any email being spam ($P(S)$) is $\frac{1}{5}$.
- The probability of the word "free" appearing in spam emails ($P(F|S)$) is $\frac{1}{2}$.
- The probability of finding the word "free" in any email ($P(F)$) is $\frac{1}{10}$.

Calculate the probability that an email is spam given that the word "free" appears in it ($P(S|F)$):

$$P(S|F) = \frac{P(F|S)P(S)}{P(F)} = \frac{\frac{1}{2} \times \frac{1}{5}}{\frac{1}{10}} = \frac{\frac{1}{10}}{\frac{1}{10}} = 1 \text{ or } 100\%$$

10 Binomial Distribution

Means that there are only two outcomes, 1 or 0, head or tails etc that are equally weighted.
To solve questions use the formula:

$$P(x = i) = \frac{\text{Probability of x = i}}{\text{Total outcomes}} \tag{10.1}$$

Finding the sum of total probabilities is found with ordered sampling with replacement:
First we find n and k:

$$\text{n = number of outcomes (always 2 for Binomial Distribution)k = total number of tests} \tag{10.2}$$
$$\tag{10.3}$$

$$n^k \tag{10.4}$$

Probability of x = i is found by using unordered sampling without replacement:
First we find n and k:

$$\text{n = total number of tests} \tag{10.5}$$
$$\text{k = number of outcomes that meet our criteria x = i} \tag{10.6}$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{10.7}$$

10.1 Examples

Coin example
If we flip a coin 5 times, what is the probability that we get 3 heads?
Find the total number of outcomes:
First we find n and k

$$n = 2(\text{possible outcomes, Head or Tail}) \tag{10.8}$$
$$k = 5(\text{Total flips}) \tag{10.9}$$
$$\tag{10.10}$$

Next we compute it:

$$2^5 = 32 \tag{10.11}$$

Next we compute the probability getting 3 heads:

$$n = 5(\text{Number of tests}) \tag{10.12}$$
$$k = 3(\text{Amount of heads required}) \tag{10.13}$$
$$\tag{10.14}$$

$$\binom{5}{3} = \frac{5!}{3!(5-3)!} \tag{10.15}$$
$$\binom{5}{3} = \frac{5!}{3!2!} \tag{10.16}$$
$$\text{You can generally reduce the expression by factoring out the second factorial} \binom{5}{3} = \frac{5*4*3}{3*2*1} \tag{10.17}$$
$$\binom{5}{3} = \frac{20}{2} \tag{10.18}$$
$$\binom{5}{3} = 10 \tag{10.19}$$
$$\tag{10.20}$$

We now have our two results, we can now compute the Binomial Distribution:

$$P(x = i) = \frac{10}{32} \tag{10.21}$$

10.2 More Complicated Example

50 students live in a dormitory. The parking lot has the capacity for 30 cars. If each student has a car with probability $\frac{1}{2}$ (independently from other students), what is the probability that there won't be enough parking spaces for all the cars.

First we compute the total number of outcomes:

$$\begin{aligned}
n &= 2 && (10.22)\\
k &= 50 && (10.23)\\
n^k = 2^{50} &= 1,125,899,906,842,624 && (10.24)\\
& && (10.25)
\end{aligned}$$

This is the total number of outcomes, now let's compute how many outcomes there are more than 30 cars:

$$\begin{aligned}
n &= 50 && (10.26)\\
k &= 31 && (10.27)\\
\binom{n}{k} &= \frac{n!}{k!(n-k)!} && (10.28)\\
\binom{50}{31} = \frac{50!}{31!(50-31)!} &= 30,405,943,383,200 && (10.29)\\
& && (10.30)
\end{aligned}$$

You then need to do this for all values from k = 31 to n and sum them. But if you just want to have 31 cars, then it can be computed like this:

$$P(x = 31) = \frac{30,405,943,383,200}{1,125,899,906,842,624} = 0.027 \tag{10.31}$$

If you want the full result you need to sum the results together:

$$\sum_{k=31}^{n} = P(x = k) \tag{10.32}$$

11 Bernoulli Distribution

The Bernoulli distribution is one of the simplest discrete distributions, involving only two possible outcomes, typically labeled as success (1) and failure (0). It is a special case of the binomial distribution with $n = 1$.

11.1 Definition

A random variable X follows a Bernoulli distribution if its probability mass function (PMF) is given by:

$$P(X = k) = \begin{cases} p & \text{if } k = 1, \\ 1 - p & \text{if } k = 0, \end{cases}$$

where p is the probability of success, $1 - p$ is the probability of failure, and k can only take values 0 or 1.

11.2 Properties

- **Mean (Expected Value):** The mean of a Bernoulli random variable X is p.
- **Variance:** The variance of X is $p(1-p)$.

11.3 Examples

Example 1: Flipping a Biased Coin
If a biased coin has a 70% chance of landing heads (success), then each flip of the coin is a Bernoulli trial where $p = 0.7$. The probability of getting tails (failure) in a single flip is $1 - 0.7 = 0.3$.

$$P(X = 1) = 0.7 \quad \text{and} \quad P(X = 0) = 0.3$$

Example 2: Product Inspection
In a factory quality control test, each product has a 95% chance of passing the inspection. Here, passing the inspection is a "success" ($k = 1$), and $p = 0.95$. The probability of a product failing the test is $1 - p = 0.05$.

$$P(X = 1) = 0.95 \quad \text{and} \quad P(X = 0) = 0.05$$

Example 3: Yes or No Survey
In a survey, respondents are asked a yes/no question where "yes" is considered a success. If historical data suggests that 40% of respondents answer "yes", then for a random respondent, the outcome of their response can be modeled as a Bernoulli random variable with $p = 0.4$.

$$P(X = 1) = 0.4 \quad \text{and} \quad P(X = 0) = 0.6$$

11.4 Conclusion

The Bernoulli distribution is widely used to model binary data and serves as the basis for more complex models involving binary or dichotomous variables, such as the logistic regression model for binary classification.

12 Probability Density Function (PDF)

PDF is used for getting the probability of each variable. It is useful to break down a problem into sub problems.
To generate a PDF you get either a list of variables or a statement.

- You then count how many total elements there are in the set of items
- Then you count many many of each variable there are
- Finally you subtract the amount of variables with the total amount of elements, do this for each PDF element

12.1 Examples

12.1.1 Throwing 2 dice

First compute the total amount of possible dice thrown using Ordered sampling with replacement:

$$n^k \tag{12.1}$$
$$n = 6(sides) \tag{12.2}$$
$$k = 2(dice) \tag{12.3}$$
$$6^2 = 36 \tag{12.4}$$
$$\tag{12.5}$$

This means there are 36 possible dice throws

12.1.2 Probability Distribution

You then count how many possible combinations of numbers there are for each and divide it with the total number.

For 4, there are 3 possible options:

$$d1 = 1, d2 = 3 \tag{12.6}$$
$$d1 = 3, d2 = 1 \tag{12.7}$$
$$d1 = 2, d2 = 2 \tag{12.8}$$
$$\tag{12.9}$$

There the probability to roll 4 is: $\frac{3}{36}$. This is then repeated for each possible number:

Sum	Number of Combinations	Probability
2	1	$\frac{1}{36}$
3	2	$\frac{2}{36} = \frac{1}{18}$
4	3	$\frac{3}{36} = \frac{1}{12}$
5	4	$\frac{4}{36} = \frac{1}{9}$
6	5	$\frac{5}{36}$
7	6	$\frac{6}{36} = \frac{1}{6}$
8	5	$\frac{5}{36}$
9	4	$\frac{4}{36} = \frac{1}{9}$
10	3	$\frac{3}{36} = \frac{1}{12}$
11	2	$\frac{2}{36} = \frac{1}{18}$
12	1	$\frac{1}{36}$

13 Cumulative Distribution Function (CDF) for Two Dice

Given the probability distribution for each possible sum when two dice are thrown, we derive the Cumulative Distribution Function (CDF), which accumulates these probabilities.

13.1 Computing the Cumulative Distribution Function (CDF)

To find the CDF at each possible sum, we accumulate the probabilities from the PDF, starting at the smallest sum, as follows:

$$\text{CDF}(x) = \sum_{i=\min}^{x} P(\text{Sum} = i)$$

where $P(\text{Sum} = i)$ is the probability for each sum from the PDF where we threw 2 dice above.

13.2 Calculated CDF Values

Sum (x)	**CDF**(x)
2	$\frac{1}{36}$
3	$\frac{1}{36} + \frac{2}{36} = \frac{3}{36}$
4	$\frac{3}{36} + \frac{3}{36} = \frac{6}{36}$
5	$\frac{6}{36} + \frac{4}{36} = \frac{10}{36}$
6	$\frac{10}{36} + \frac{5}{36} = \frac{15}{36}$
7	$\frac{15}{36} + \frac{6}{36} = \frac{21}{36}$
8	$\frac{21}{36} + \frac{5}{36} = \frac{26}{36}$
9	$\frac{26}{36} + \frac{4}{36} = \frac{30}{36}$
10	$\frac{30}{36} + \frac{3}{36} = \frac{33}{36}$
11	$\frac{33}{36} + \frac{2}{36} = \frac{35}{36}$
12	$\frac{35}{36} + \frac{1}{36} = 1$

Each CDF value represents the probability of rolling a sum that is less than or equal to x. For instance, the CDF for the sum of 4 indicates that there is a $\frac{1}{6}$ chance of rolling at least 4 with two dice.

14 Expected Value and Variance

The Expected Value (mean) and Variance are fundamental concepts in statistics that provide key insights into the behavior of probability distributions.

14.1 Expected Value (Mean)

The expected value of a random variable gives a measure of the center of the distribution of the variable. It's defined as the weighted average of all possible values that this random variable can take on, with the weights being the probabilities of the respective values.

$$E(X) = \sum_{i=1}^{n} x_i P(x_i)$$

Where:

- x_i are the values the random variable X can take.
- $P(x_i)$ are the probabilities associated with each value.

14.2 Variance

Variance measures the spread of a set of values. It's calculated by taking the average of the squared differences from the Mean.

$$\text{Var}(X) = E[(X - \mu)^2] = \sum_{i=1}^{n} (x_i - \mu)^2 P(x_i)$$

Where:

- μ is the Expected Value of X.

Simplified version:

$$\text{VAR(X)} = E(X^2) \text{ - } |E(X)|^2$$

14.3 Varience example from 2022 exam

Given PDF:

$$\frac{1}{4}x = -2 \tag{14.1}$$
$$\frac{1}{4}x = 0 \tag{14.2}$$
$$\frac{1}{4}x = 1 \tag{14.3}$$
$$\frac{1}{4}x = 3 \tag{14.4}$$
$$\tag{14.5}$$

First compute E(x):

$$E(X) = -2 \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} = \frac{14}{4} \tag{14.6}$$

Next compute $E(X^2)$:

$$E(X^2) = -2^2 \cdot \frac{1}{4} + 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{1}{4} + 3^2 \cdot \frac{1}{4} = \frac{14}{4} \tag{14.7}$$

For:

$$E(X) = \frac{2}{4} \tag{14.8}$$
$$E(X^2) = \frac{14}{4} \tag{14.9}$$
$$VAR(X) = \frac{14}{4} - (\frac{2}{4})^2 \tag{14.10}$$
$$VAR(X) = \frac{14}{4} - \frac{4}{16} \tag{14.11}$$
$$VAR(X) = \frac{14}{4} - \frac{1}{4} \tag{14.12}$$
$$VAR(X) = \frac{13}{4} \tag{14.13}$$
$$\tag{14.14}$$

14.4 Calculations for Expected Value and Variance

The expected value and variance are crucial statistical measures that help understand the central tendency and variability of a probability distribution.

14.4.1 Expected Value (E(X))

To compute the expected value $E(X)$, sum up the products of each value the variable X can take and its corresponding probability from the PDF.

$$E(X) = \sum x_i P(x_i)$$

Given the PDF for dice, the calculation would proceed as follows:

$$E(X) = 0 \cdot \frac{1}{34} + 2 \cdot \frac{2}{34} + 4 \cdot \frac{13}{34} + 7 \cdot \frac{9}{34} + 10 \cdot \frac{8}{34} + 12 \cdot \frac{1}{34} \tag{14.15}$$
$$= \frac{0}{34} + \frac{4}{34} + \frac{52}{34} + \frac{63}{34} + \frac{80}{34} + \frac{12}{34} \tag{14.16}$$
$$= \frac{211}{34} \tag{14.17}$$
$$= 6.21 \tag{14.18}$$

14.4.2 Variance of X (Var(X))

To calculate the variance $\text{Var}(X)$, first compute $E(X^2)$, the expected value of the square of X.

$$E(X^2) = \sum x_i^2 P(x_i)$$

$$\begin{aligned}
E(X^2) &= 0^2 \cdot \frac{1}{34} + 2^2 \cdot \frac{2}{34} + 4^2 \cdot \frac{13}{34} + 7^2 \cdot \frac{9}{34} + 10^2 \cdot \frac{8}{34} + 12^2 \cdot \frac{1}{34} && (14.19)\\
&= \frac{0}{34} + \frac{8}{34} + \frac{208}{34} + \frac{441}{34} + \frac{800}{34} + \frac{144}{34} && (14.20)\\
&= \frac{1601}{34} && (14.21)\\
&= 47.09 && (14.22)
\end{aligned}$$

Now, compute the variance:

$$\begin{aligned}
\text{Var}(X) &= E(X^2) - [E(X)]^2 && (14.23)\\
&= 47.09 - (6.21)^2 && (14.24)\\
&= 47.09 - 38.56 && (14.25)\\
&= 8.53 && (14.26)
\end{aligned}$$

14.5 PMF of grades from Assignment 5

To start with we count how many total grades there are listed, which is 34. Then we can simply count each grade and create the PMF:

$$\begin{aligned}
&0 \text{ if x = -3} && (14.27)\\
&\frac{1}{34} \text{ if x = 0} && (14.28)\\
&\frac{2}{34} \text{ if x = 2} && (14.29)\\
&\frac{13}{34} \text{ if x = 4} && (14.30)\\
&\frac{9}{34} \text{ if x = 7} && (14.31)\\
&\frac{8}{34} \text{ if x = 10} && (14.32)\\
&\frac{1}{34} \text{ if x = 12} && (14.33)\\
& && (14.34)
\end{aligned}$$

If we now sum them together then it should equal 1:

$$\frac{1}{34} + \frac{2}{34} + \frac{13}{34} + \frac{9}{34} + \frac{8}{34} + \frac{1}{34} = \frac{34}{34} = 1 \tag{14.35}$$

This then confirms that our PMF is configured correctly.

14.5.1 Expected x value

To compute E(x) we simply sum the PDF:

$$\begin{aligned}
E(x) = 0 \cdot \frac{1}{34} + 2 \cdot \frac{2}{34} + 4 \cdot \frac{13}{34} + 7 \cdot \frac{9}{34} + 10 \cdot \frac{8}{34} + 12 \cdot \frac{1}{34} && (14.36)\\
E(x) = \frac{4}{34} + \frac{52}{34} + \frac{63}{34} + \frac{80}{34} + \frac{12}{34} && (14.37)\\
E(x) = \frac{211}{34} && (14.38)\\
E(x) = 6.21 && (14.39)
\end{aligned}$$

14.5.2 Variance of x

To calculate var(x) we first need to compute $E(x^2)$:

$$E(x) = \sum x_i P(x_i) \tag{14.40}$$
$$E(x^2) = 0^2 \cdot \frac{1}{34} + 2^2 \cdot \frac{2}{34} + 4^2 \cdot \frac{13}{34} + 7^2 \cdot \frac{9}{34} + 10^2 \cdot \frac{8}{34} + 12^2 \cdot \frac{1}{34} \tag{14.41}$$
$$E(x^2) = 4 \cdot \frac{2}{34} + 16 \cdot \frac{13}{34} + 49 \cdot \frac{9}{34} + 100 \cdot \frac{8}{34} + 144 \cdot \frac{1}{34} \tag{14.42}$$
$$E(x^2) = 47.09 \tag{14.43}$$

We can now compute the varience:

$$\text{Var}(x) = E(x^2) - [E(x)]^2 var(x) = 47.09 - 6.21^2 var(x) = 8.53 \tag{14.44}$$

14.5.3 CDF

Finally we compute the CDF by summing each element of the PDF with the CDF of the previous element:

$$\text{0 if x <= -3} \tag{14.45}$$
$$\frac{1}{34} + 0 = \frac{1}{34} \text{ if x <= 0} \tag{14.46}$$
$$\frac{2}{34} + \frac{1}{34} = \frac{3}{34} \text{ if x <= 2} \tag{14.47}$$
$$\frac{13}{34} + \frac{3}{34} = \frac{16}{34} \text{ if x <= 4} \tag{14.48}$$
$$\frac{9}{34} + \frac{16}{34} = \frac{25}{34} \text{ if x <= 7} \tag{14.49}$$
$$\frac{8}{34} + \frac{25}{34} = \frac{33}{34} \text{ if x <= 10} \tag{14.50}$$
$$\frac{1}{34} + \frac{33}{34} = \frac{34}{34} = 1 \text{ if x <= 12} \tag{14.51}$$
$$\tag{14.52}$$

15 Joint Probability

16 Joint Probability Distributions

Joint probability distributions are useful for studying the relationship between two random variables. This section covers the joint probability mass function (PMF), the joint cumulative distribution function (CDF), conditioning and independence, conditional expectation and variance, and covariance and correlation.

16.1 Joint Probability Mass Function (PMF) Example

Consider two random variables X and Y, each representing the outcome of a die roll (dice are fair and six-sided).

16.1.1 Computing the Joint PMF

The joint PMF $P(X = x, Y = y)$ is the probability that X equals x and Y equals y. Because the dice rolls are independent, the joint PMF is the product of their individual probabilities.

$$P(X = x, Y = y) = P(X = x) \times P(Y = y) = \frac{1}{6} \times \frac{1}{6} = \frac{1}{36}$$

This holds for each pair (x, y) where $x, y \in \{1, 2, 3, 4, 5, 6\}$.

16.1.2 Complete Joint PMF Table

Here is the complete table for the joint PMF of two dice:

X\Y	**1**	**2**	**3**	**4**	**5**	**6**
1	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

16.2 Marginal Probability Mass Function (PMF)

The marginal PMF of a variable is found by summing the joint probabilities over all values of the other variable.

$$P_X(x) = \sum_{y \in Y} P(X = x, Y = y)$$

$$P_Y(y) = \sum_{x \in X} P(X = x, Y = y)$$

16.2.1 Example 1: Marginal PMF of X

Given the joint probability mass function (PMF) table:

X\Y	1	2
1	0.5	0.1
2	0.3	0.1

To find the marginal PMF of X, we sum the probabilities across each row:

- For $X = 1$, sum the probabilities in the first row:

$$P_X(1) = 0.5 + 0.1 = 0.6$$

- For $X = 2$, sum the probabilities in the second row:

$$P_X(2) = 0.3 + 0.1 = 0.4$$

The marginal PMF of X is thus $P_X(1) = 0.6$ and $P_X(2) = 0.4$.

16.2.2 Example 2: Marginal PMF of Y

For the same table, to find the marginal PMF of Y, sum the probabilities down each column:

- For $Y = 1$, sum the probabilities in the first column:

$$P_Y(1) = 0.5 + 0.3 = 0.8$$

- For $Y = 2$, sum the probabilities in the second column:

$$P_Y(2) = 0.1 + 0.1 = 0.2$$

The marginal PMF of Y is thus $P_Y(1) = 0.8$ and $P_Y(2) = 0.2$.

16.3 Conditional Probability Mass Function

The conditional PMF describes the probability of one event given the occurrence of another:

$$P(Y = y \mid X = x) = \frac{P(X = x, Y = y)}{P_X(x)}$$

16.3.1 Example 1: Conditional PMF from Joint PMF

Using the joint PMF table, to find $P(Y = 2 \mid X = 1)$, follow these steps:

- Identify the joint probability $P(X = 1, Y = 2)$:

$$P(X = 1, Y = 2) = 0.1$$

- Find the marginal probability $P_X(1)$, which is the sum of the probabilities where $X = 1$:

$$P_X(1) = 0.5 + 0.1 = 0.6$$

- Calculate the conditional probability using the formula:

$$P(Y = 2 \mid X = 1) = \frac{0.1}{0.6} = \frac{1}{6}$$

This calculation shows that the probability of $Y = 2$ given $X = 1$ is $\frac{1}{6}$.

16.3.2 Example 2: Another Conditional PMF

To determine $P(Y = 1 \mid X = 2)$ from the same table:

- Identify the joint probability $P(X = 2, Y = 1)$:

$$P(X = 2, Y = 1) = 0.3$$

- Find the marginal probability $P_X(2)$, which is the sum of the probabilities where $X = 2$:

$$P_X(2) = 0.3 + 0.1 = 0.4$$

- Apply the formula for conditional probability:

$$P(Y = 1 \mid X = 2) = \frac{0.3}{0.4} = 0.75$$

This indicates that the probability of $Y = 1$ given $X = 2$ is 0.75, or 75%.

16.4 Conditional Probability for Events

This section covers how to find the probability of an event given that another event has occurred:

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

16.5 Conditional Probability for Events

This section covers how to find the probability of an event given that another event has occurred:

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

16.5.1 Example Using the Provided Table

Suppose in the context of the given PMF table:

X\Y	1	2
1	0.5	0.1
2	0.3	0.1

we define event A as $X = 1$ and event B as $Y = 1$. Then, $P(A \cap B)$ represents the joint probability of $X = 1$ and $Y = 1$, which from the table is:

$$P(A \cap B) = 0.5$$

If the probability of B (i.e., $P(B)$) is the sum of probabilities where $Y = 1$ across all X, then:

$$P(B) = 0.5 + 0.3 = 0.8$$

Applying the formula for conditional probability:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.5}{0.8} = 0.625$$

This calculation demonstrates a 62.5% chance that $X = 1$ given $Y = 1$.

16.5.2 Example 1: Probability of Raining Given Carrying an Umbrella

If the probability of raining (A) and carrying an umbrella (B) are correlated such that $P(A \cap B) = 0.2$ and $P(B) = 0.5$, then calculate the conditional probability:

- Determine the probability of both events happening together ($P(A \cap B)$):

$$P(A \cap B) = 0.2$$

- Determine the probability of the conditioning event ($P(B)$):

$$P(B) = 0.5$$

- Apply the formula for conditional probability:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0.2}{0.5} = 0.4$$

This means there is a 40% chance of raining given that an individual is carrying an umbrella.

16.5.3 Example 2: Probability of Liking Chocolate Given Liking Vanilla

Consider a survey result where 30% of respondents like both chocolate and vanilla ice cream, and 60% like vanilla:

- Determine the joint probability of liking both flavors ($P(\text{Chocolate} \cap \text{Vanilla})$):

$$P(\text{Chocolate} \cap \text{Vanilla}) = 0.3$$

- Determine the probability of the conditioning event ($P(\text{Vanilla})$):

$$P(\text{Vanilla}) = 0.6$$

- Apply the formula for conditional probability:

$$P(\text{Chocolate} \mid \text{Vanilla}) = \frac{P(\text{Chocolate} \cap \text{Vanilla})}{P(\text{Vanilla})} = \frac{0.3}{0.6} = 0.5$$

This indicates that there is a 50% chance of a person liking chocolate given that they like vanilla ice cream.

16.6 Checking Independence of Variables

Two variables are independent if the joint distribution is the product of their marginal distributions:

$$X \text{ and } Y \text{ are independent if } P(X = x, Y = y) = P_X(x) \cdot P_Y(y) \text{ for all } x \text{ and } y.$$

16.6.1 Example 1: Checking Independence from Joint PMF

X\Y	1	2
1	0.5	0.1
2	0.3	0.1

To determine if variables X and Y from our joint PMF table are independent:

- Observe the joint probability $P(X = 1, Y = 1)$:

$$P(X = 1, Y = 1) = 0.5$$

- Calculate the product of the marginal probabilities for $X = 1$ and $Y = 1$:
  - First, identify the marginal probabilities from the provided data:
    * $P_X(1)$ is the sum of all probabilities in the row of the joint PMF table where $X = 1$. This is calculated as follows:

$$P_X(1) = P(X = 1, Y = 1) + P(X = 1, Y = 2) = 0.5 + 0.1 = 0.6$$

    * $P_Y(1)$ is the sum of all probabilities in the column of the joint PMF table where $Y = 1$. This is calculated as follows:

$$P_Y(1) = P(X = 1, Y = 1) + P(X = 2, Y = 1) = 0.5 + 0.3 = 0.8$$

- Compute the product $P_X(1) \cdot P_Y(1)$:

$$P_X(1) \cdot P_Y(1) = 0.6 \times 0.8 = 0.48$$

- Compare the joint probability with the product of the marginal probabilities:

$$P(X = 1, Y = 1) = 0.5 \neq 0.48$$

Since $P(X = 1, Y = 1)$ is not equal to $P_X(1) \cdot P_Y(1)$, X and Y are not independent.

16.6.2 Example 2: Independence in Another Setup

For a different set of probabilities:

- Observe the joint probability $P(X = 1, Y = 1)$:

$$P(X = 1, Y = 1) = 0.24$$

- Calculate the marginal probabilities and their product:

$$P_X(1) = 0.4 \quad \text{and} \quad P_Y(1) = 0.6$$

- Compute the product $P_X(1) \cdot P_Y(1)$:

$$P_X(1) \cdot P_Y(1) = 0.4 \times 0.6 = 0.24$$

- Compare the joint probability with the product of the marginal probabilities:

$$P(X = 1, Y = 1) = 0.24 = 0.24$$

Since $P(X = 1, Y = 1)$ equals $P_X(1) \cdot P_Y(1)$, X and Y are independent in this case.

16.7 Expected Value Given a Condition

The expected value of a random variable given a condition is the sum of the products of all possible values of the variable and their respective conditional probabilities.

$$E[X \mid Y = y] = \sum_x x \cdot P(X = x \mid Y = y)$$

16.7.1 Example 1

Given the joint PMF table:

X\Y	1	2
1	0.5	0.1
2	0.3	0.1

To calculate $E[X \mid Y = 1]$:

- First, determine the conditional probabilities $P(X = x \mid Y = 1)$ for each X:

$$P(X = 1 \mid Y = 1) = \frac{0.5}{0.5 + 0.3} = 0.625$$

$$P(X = 2 \mid Y = 1) = \frac{0.3}{0.5 + 0.3} = 0.375$$

- Then, calculate the expected value as the sum of the products of each X value and its conditional probability:

$$E[X \mid Y = 1] = 1 \cdot 0.625 + 2 \cdot 0.375 = 0.625 + 0.75 = 1.375$$

The calculation was adjusted for accuracy based on the probabilities computed from the table.

16.7.2 Example 2

Assuming a scenario with given probabilities:

- Determine the conditional probabilities $P(X = x \mid Y = 2)$ for each X:

$$P(X = 1 \mid Y = 2) = \frac{0.1}{0.1 + 0.1} = 0.5$$

$$P(X = 2 \mid Y = 2) = \frac{0.1}{0.1 + 0.1} = 0.5$$

- Calculate the expected value:

$$E[X \mid Y = 2] = 1 \cdot 0.5 + 2 \cdot 0.5 = 0.5 + 1 = 1.5$$

This results in an expected value of 1.5 for X given $Y = 2$, adjusted from the original miscalculation.

16.8 Variance Given a Condition

The variance of a random variable given a condition is calculated using the formula:

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

where

$$E[X^2 \mid Y = y] = \sum_x x^2 \cdot P(X = x \mid Y = y)$$

16.8.1 Example 1

First, calculate $E[X^2 \mid Y = 1]$ using the conditional probabilities derived previously:

$$E[X^2 \mid Y = 1] = 1^2 \cdot 0.625 + 2^2 \cdot 0.375 = 0.625 + 1.5 = 2.125$$

Then, using the previously calculated $E[X \mid Y = 1] = 1.375$:

$$\text{Var}(X \mid Y = 1) = 2.125 - (1.375)^2 = 2.125 - 1.8906 = 0.2344$$

This calculation indicates that the variance of X given $Y = 1$ is 0.2344.

16.8.2 Example 2

Assume we know $E[X^2 \mid Y = 2]$ and $E[X \mid Y = 2]$ from previous calculations:

$$E[X^2 \mid Y = 2] = 1^2 \cdot 0.5 + 2^2 \cdot 0.5 = 0.5 + 2 = 2.5$$

And using $E[X \mid Y = 2] = 1.5$:

$$\text{Var}(X \mid Y = 2) = 2.5 - (1.5)^2 = 2.5 - 2.25 = 0.25$$

Thus, the variance of X given $Y = 2$ is 0.25.

16.8.3 Computing the Joint CDF

To find $F(3, 4)$, for instance, we need to sum all probabilities where $X \leq 3$ and $Y \leq 4$.

$$F(3, 4) = \sum_{x=1}^{3} \sum_{y=1}^{4} P(X = x, Y = y) = 12 \times \frac{1}{36} = \frac{12}{36} = \frac{1}{3}$$

16.8.4 Partial Joint CDF Table

Below is a part of the joint CDF table up to $x, y = 3$:

X\Y	**1**	**2**	**3**
1	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$
2	$\frac{2}{36}$	$\frac{4}{36}$	$\frac{6}{36}$
3	$\frac{3}{36}$	$\frac{6}{36}$	$\frac{9}{36}$

16.8.5 Explanation

Each entry in the CDF table is the cumulative sum of the probabilities from the joint PMF table where x and y values do not exceed the row and column indices, respectively.

16.9 Example of working with Joint PMF

Given the joint PMF of random variables X and Y:

X\Y	**1**	**2**
1	$\frac{1}{3}$	$\frac{1}{12}$
2	$\frac{1}{6}$	0
4	$\frac{1}{12}$	$\frac{1}{3}$

16.9.1 Find $P(X \leq 2, Y > 1)$

To find this probability, we sum the probabilities from the joint PMF table where $X \leq 2$ and $Y > 1$.

$$P(X \leq 2, Y > 1) = P(X = 1, Y = 2) + P(X = 2, Y = 2) = \frac{1}{12} + 0 = \frac{1}{12}$$

16.9.2 Find the marginal PMFs of X and Y

The marginal PMF of X, denoted $P_X(x)$, is found by summing over Y:

$$P_X(1) = P(X = 1, Y = 1) + P(X = 1, Y = 2) = \frac{1}{3} + \frac{1}{12} = \frac{5}{12}$$

$$P_X(2) = P(X = 2, Y = 1) + P(X = 2, Y = 2) = \frac{1}{6} + 0 = \frac{1}{6}$$

$$P_X(4) = P(X = 4, Y = 1) + P(X = 4, Y = 2) = \frac{1}{12} + \frac{1}{3} = \frac{5}{12}$$

The marginal PMF of Y, denoted $P_Y(y)$, is found by summing over X:

$$P_Y(1) = P(X = 1, Y = 1) + P(X = 2, Y = 1) + P(X = 4, Y = 1) = \frac{1}{3} + \frac{1}{6} + \frac{1}{12} = \frac{1}{2}$$

$$P_Y(2) = P(X = 1, Y = 2) + P(X = 2, Y = 2) + P(X = 4, Y = 2) = \frac{1}{12} + 0 + \frac{1}{3} = \frac{5}{12}$$

16.9.3 Find $P(Y = 2 \mid X = 1)$

This is the probability of $Y = 2$ given that $X = 1$. We use the formula for conditional probability:

$$P(Y = 2 \mid X = 1) = \frac{P(X = 1, Y = 2)}{P_X(1)} = \frac{\frac{1}{12}}{\frac{5}{12}} = \frac{1}{5}$$

16.9.4 4. Are X and Y independent?

To check for independence, we compare $P(X = x, Y = y)$ to $P_X(x) \times P_Y(y)$ for all x, y.

$$P(X = 1, Y = 1) = \frac{1}{3}, \quad P_X(1) \times P_Y(1) = \frac{5}{12} \times \frac{1}{2} = \frac{5}{24}$$

Since $\frac{1}{3} \neq \frac{5}{24}$, X and Y are not independent.

16.10 computing new variable Z

Let X and Y be as defined in Problem 1, and let a new random variable $Z = X - 2Y$. We need to find the probability mass function (PMF) of Z and calculate $P(X = 2 \mid Z = 0)$.

X\Y	1	2
1	$\frac{1}{3}$	$\frac{1}{12}$
2	$\frac{1}{6}$	0
4	$\frac{1}{12}$	$\frac{1}{3}$

(16.1)

16.10.1 Find PMF of Z

To find the PMF of Z, we first calculate all possible values of Z given the joint occurrences of X and Y, and then aggregate the probabilities.
The possible pairs (X, Y) with their corresponding Z values and probabilities are:

$$(X = 1, Y = 1) \Rightarrow Z = 1 - 2 \cdot 1 = -1, \quad P(Z = -1) = \frac{1}{3} \text{Found by looking at } (X = 1, Y = 1) \text{in PMF}$$
$$(X = 1, Y = 2) \Rightarrow Z = 1 - 2 \cdot 2 = -3, \quad P(Z = -3) = \frac{1}{12}$$
$$(X = 2, Y = 1) \Rightarrow Z = 2 - 2 \cdot 1 = 0, \quad P(Z = 0) = \frac{1}{6}$$
$$(X = 4, Y = 1) \Rightarrow Z = 4 - 2 \cdot 1 = 2, \quad P(Z = 2) = \frac{1}{12}$$
$$(X = 4, Y = 2) \Rightarrow Z = 4 - 2 \cdot 2 = 0, \quad P(Z = 0) = \frac{1}{3}$$

Summing the probabilities for $Z = 0$ since there are two scenarioes:

$$P(Z = 0) = \frac{1}{6} + \frac{1}{3} = \frac{1}{2}$$

Thus, the PMF of Z is:

Z	Probability
-3	$\frac{1}{12}$
-1	$\frac{1}{3}$
0	$\frac{1}{2}$
2	$\frac{1}{12}$

16.10.2 Find $P(X = 2 \mid Z = 0)$

Using the conditional probability formula:

$$P(X = 2 \mid Z = 0) = \frac{P(X = 2 \text{ and } Z = 0)}{P(Z = 0)}$$

We already know from the joint events:

$$P(X = 2 \text{ and } Z = 0) = \frac{1}{6}, \quad P(Z = 0) = \frac{1}{2}$$

Thus,

$$P(X = 2 \mid Z = 0) = \frac{\frac{1}{6}}{\frac{1}{2}} = \frac{1}{3}$$

16.11 Joint Cumulative Distribution Function (CDF) Example

The joint CDF $F(x, y)$ is defined as the probability that $X \leq x$ and $Y \leq y$.

16.12 Conditional PMF, Expectation and varience example

Given that X and Y are defined as in Problem 1, and it is given that $Y = 1$, we are to solve the following:

16.12.1 Find the conditional PMF of X given $Y = 1$

To find the conditional PMF $P_{X|Y}(x|1)$, we use:

$$P_{X|Y}(x|1) = \frac{P(X = x, Y = 1)}{P(Y = 1)}$$

Given from the joint PMF table:

$$\begin{aligned} P(Y = 1) &= P(X = 1, Y = 1) + P(X = 2, Y = 1) + P(X = 4, Y = 1) \\ &= \frac{1}{3} + \frac{1}{6} + \frac{1}{12} = \frac{7}{12} \end{aligned}$$

Calculating for each x:

$$P_{X|Y}(1|1) = \frac{\frac{1}{3}}{\frac{7}{12}} = \frac{4}{7}, \quad P_{X|Y}(2|1) = \frac{\frac{1}{6}}{\frac{7}{12}} = \frac{2}{7}, \quad P_{X|Y}(4|1) = \frac{\frac{1}{12}}{\frac{7}{12}} = \frac{1}{7}$$

16.12.2 Find $E[X|Y = 1]$

The expected value of X given $Y = 1$ is calculated as:

$$E[X|Y = 1] = \sum_x x \cdot P_{X|Y}(x|1)$$

$$E[X|Y = 1] = 1 \cdot \frac{4}{7} + 2 \cdot \frac{2}{7} + 4 \cdot \frac{1}{7} = \frac{4}{7} + \frac{4}{7} + \frac{4}{7} = \frac{12}{7}$$

16.12.3 Find $\text{Var}(X|Y = 1)$

The variance given $Y = 1$ is computed using:

$$\text{Var}(X|Y = 1) = E[X^2|Y = 1] - (E[X|Y = 1])^2$$

Calculating $E[X^2|Y = 1]$:

$$E[X^2|Y = 1] = 1^2 \cdot \frac{4}{7} + 2^2 \cdot \frac{2}{7} + 4^2 \cdot \frac{1}{7} = \frac{4}{7} + \frac{8}{7} + \frac{16}{7} = \frac{28}{7} = 4$$

Thus,

$$\text{Var}(X|Y = 1) = 4 - \left(\frac{12}{7}\right)^2 = 4 - \frac{144}{49} = \frac{52}{49}$$

16.13 Exam example with multiple independent variables

Sarah reads Y books a month, where Y can take values 0, 1, 2, or 3 with given probabilities. David reads $X = 3Y + Z$ books in the same month, where Z is the number of books he borrows, taking values 0 or 2 with given probabilities. We need to find the expectation and variance of X.

16.13.1 Given Probabilities

$$P(Y=0)=\frac{1}{10},\quad P(Y=1)=\frac{3}{10},$$
$$P(Y=2)=\frac{2}{5},\quad P(Y=3)=\frac{1}{5},$$
$$P(Z=0)=\frac{3}{5},\quad P(Z=2)=\frac{2}{5}.$$

16.13.2 Expectation and Variance Calculations

Since $X=3Y+Z$ and assuming Y and Z are independent, we calculate the expectation and variance as follows:

16.13.3 Expectation

The expectation of X is:

$$E[X]=E[3Y+Z]=3E[Y]+E[Z]$$

We calculate $E[Y]$ and $E[Z]$ using the definitions of expectation for discrete random variables:

$$E[Y]=0\cdot\frac{1}{10}+1\cdot\frac{3}{10}+2\cdot\frac{2}{5}+3\cdot\frac{1}{5}=0+0.3+0.8+0.6=1.7$$

$$E[Z]=0\cdot\frac{3}{5}+2\cdot\frac{2}{5}=0+0.8=0.8$$

Thus,

$$E[X]=3\cdot 1.7+0.8=5.1+0.8=5.9$$

16.13.4 Variance

The variance of X is:

$$\mathrm{Var}(X)=\mathrm{Var}(3Y+Z)=9\mathrm{Var}(Y)+\mathrm{Var}(Z)$$

Note constants are always scaled by C^2 because We calculate $\mathrm{Var}(Y)$ and $\mathrm{Var}(Z)$ using the formulas:

$$\mathrm{Var}(Y)=E[Y^2]-(E[Y])^2$$
$$E[Y^2]=0^2\cdot\frac{1}{10}+1^2\cdot\frac{3}{10}+2^2\cdot\frac{2}{5}+3^2\cdot\frac{1}{5}=0+0.3+1.6+1.8=3.7$$
$$\mathrm{Var}(Y)=3.7-(1.7)^2=3.7-2.89=0.81$$
$$\mathrm{Var}(Z)=E[Z^2]-(E[Z])^2$$
$$E[Z^2]=0^2\cdot\frac{3}{5}+2^2\cdot\frac{2}{5}=0+1.6=1.6$$
$$\mathrm{Var}(Z)=1.6-(0.8)^2=1.6-0.64=0.96$$

Thus,

$$\mathrm{Var}(X)=9\cdot 0.81+0.96=7.29+0.96=8.25$$

16.13.5 Explanation of Variance Scaling

The term $9\mathrm{Var}(Y)$ comes from the property of variance where scaling a random variable aY by a factor a results in the variance being scaled by a^2:

$$\mathrm{Var}(aY)=a^2\mathrm{Var}(Y)$$

For $a=3$, $a^2=9$, hence $\mathrm{Var}(3Y)=9\mathrm{Var}(Y)$.

17 Fundamental Theorem of Calculus

The Fundamental Theorem of Calculus connects differentiation and integration, showing that these two operations are essentially inverses of each other. It is divided into two parts:

17.1 Part 1

This part establishes that the definite integral of a function can be reversed by differentiation:

$$\text{If } F(x) = \int_a^x f(t)\,dt, \text{ then } F'(x) = f(x).$$

This means that if F is defined as the antiderivative of f from a to x, then the derivative of F is $f(x)$.

17.2 Part 2

This part provides a way to evaluate definite integrals:

$$\text{If } F \text{ is any antiderivative of } f, \text{ then } \int_a^b f(x)\,dx = F(b) - F(a).$$

This states that the integral of f from a to b is the difference between the values of any antiderivative F of f at the upper and lower limits of integration.
These principles not only simplify the computation of integrals but also provide a profound connection between the integral and derivative of a function.

18 Basic Rules

18.1 Quadratic Formula

The solutions to the quadratic equation $ax^2 + bx + c = 0$ are given by:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

18.2 Properties of Exponents

- $a^m \cdot a^n = a^{m+n}$
- $\frac{a^m}{a^n} = a^{m-n}$
- $(a^m)^n = a^{mn}$
- $a^{-n} = \frac{1}{a^n}$ (Inversion of exponents)
- $a^{\frac{1}{n}} = \sqrt[n]{a}$ (Radicals as exponents)

18.3 Fraction Manipulations

- Addition: $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{bd}$
- Subtraction: $\frac{a}{b} - \frac{c}{d} = \frac{ad-bc}{bd}$
- Multiplication: $\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}$
- Division: $\frac{\frac{a}{b}}{\frac{c}{d}} = \frac{a}{b} \cdot \frac{d}{c}$
- Inverting a fraction: $\frac{a}{b}^{-1} = \frac{b}{a}$

18.4 Multiplying Expressions

- **Single Terms and Parentheses:** To multiply a single term by a polynomial in parentheses, distribute the term across each term inside the parenthesis. For example:

$$2 \times (3x + 4) = 2 \cdot 3x + 2 \cdot 4 = 6x + 8$$

- **Binomial Products (FOIL Method):** To multiply two binomials, use the FOIL method:

$$(a + b)(c + d) = ac + ad + bc + bd$$

Example:

$$(x + 2)(x - 3) = x^2 - 3x + 2x - 6 = x^2 - x - 6$$

18.5 Special Polynomial Products

- **Difference of Squares:**

$$(a + b)(a - b) = a^2 - b^2$$

- **Perfect Square Trinomials:**

$$(a + b)^2 = a^2 + 2ab + b^2$$
$$(a - b)^2 = a^2 - 2ab + b^2$$

18.6 Powers of a Binomial

- **Square of a Binomial:**

$$(a + b)^2 = a^2 + 2ab + b^2$$

- **Cube of a Binomial:**

$$(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

18.7 Properties of Radicals

- $\sqrt{a} \cdot \sqrt{b} = \sqrt{ab}$
- $\frac{\sqrt{a}}{\sqrt{b}} = \sqrt{\frac{a}{b}}$
- $\sqrt[n]{a^m} = a^{\frac{m}{n}}$
- $\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{ab}$

18.8 Properties of Logarithms

- $\log_b(xy) = \log_b x + \log_b y$
- $\log_b\left(\frac{x}{y}\right) = \log_b x - \log_b y$
- $\log_b(x^n) = n \log_b x$
- Change of base formula: $\log_b x = \frac{\log_k x}{\log_k b}$

19 Geometry

19.1 Area and Perimeter

- Rectangle: Area $= ab$, Perimeter $= 2(a + b)$
- Circle: Area $= \pi r^2$, Circumference $= 2\pi r$

19.2 Pythagorean Theorem

- In a right triangle: $a^2 + b^2 = c^2$

19.3 ln rewrite rules

Quotient Rule: $\ln\left(\frac{x}{y}\right) = \ln(x) - \ln(y)$
This rule helps you rewrite the logarithm of a fraction as a difference.

Power Rule: $\ln(x^n) = n \cdot \ln(x)$
You can bring exponents in an argument out in front, which is helpful for simplifying differentiation and integration.

Square Roots: $\ln(\sqrt{x}) = \ln\big(x^{1/2}\big) = \frac{1}{2}\ln(x)$
For expressions under square roots, rewrite them with a fractional exponent and apply the power rule.