Computer Science

CSC196: Analyzing Data

Continuous Random Variables

Jason Pacheco and Cesim Erten

Outline

- Concepts of Calculus
- Continuous Probability Distributions
- Fundamental Rules of Probability

Outline

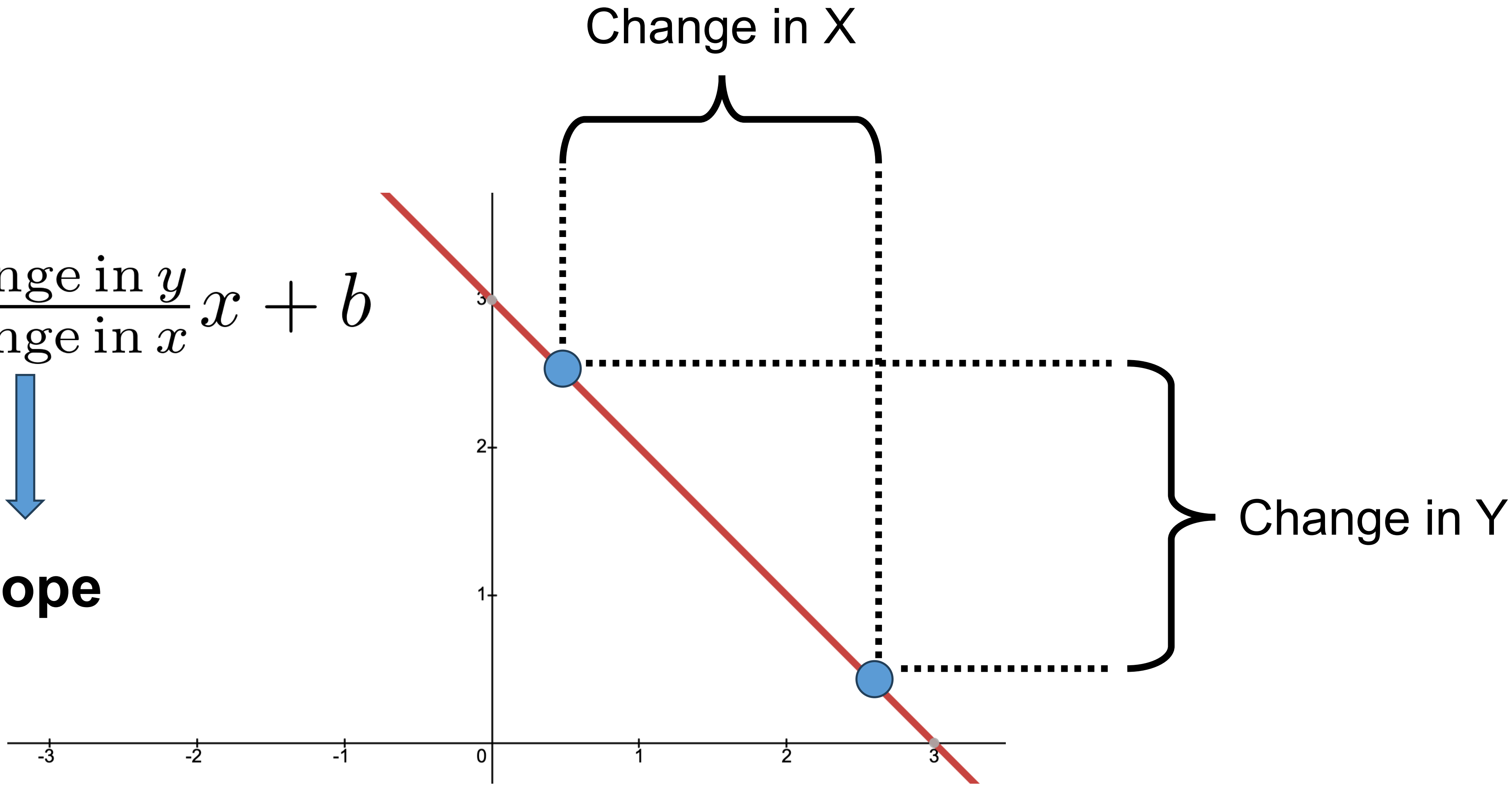
- Concepts of Calculus
- Continuous Probability Distributions
- Fundamental Rules of Probability

Equation for a Line

Recall the equation for a line in point / slope form…

$$y = \frac{\text{Change in } y}{\text{Change in } x} x + b$$

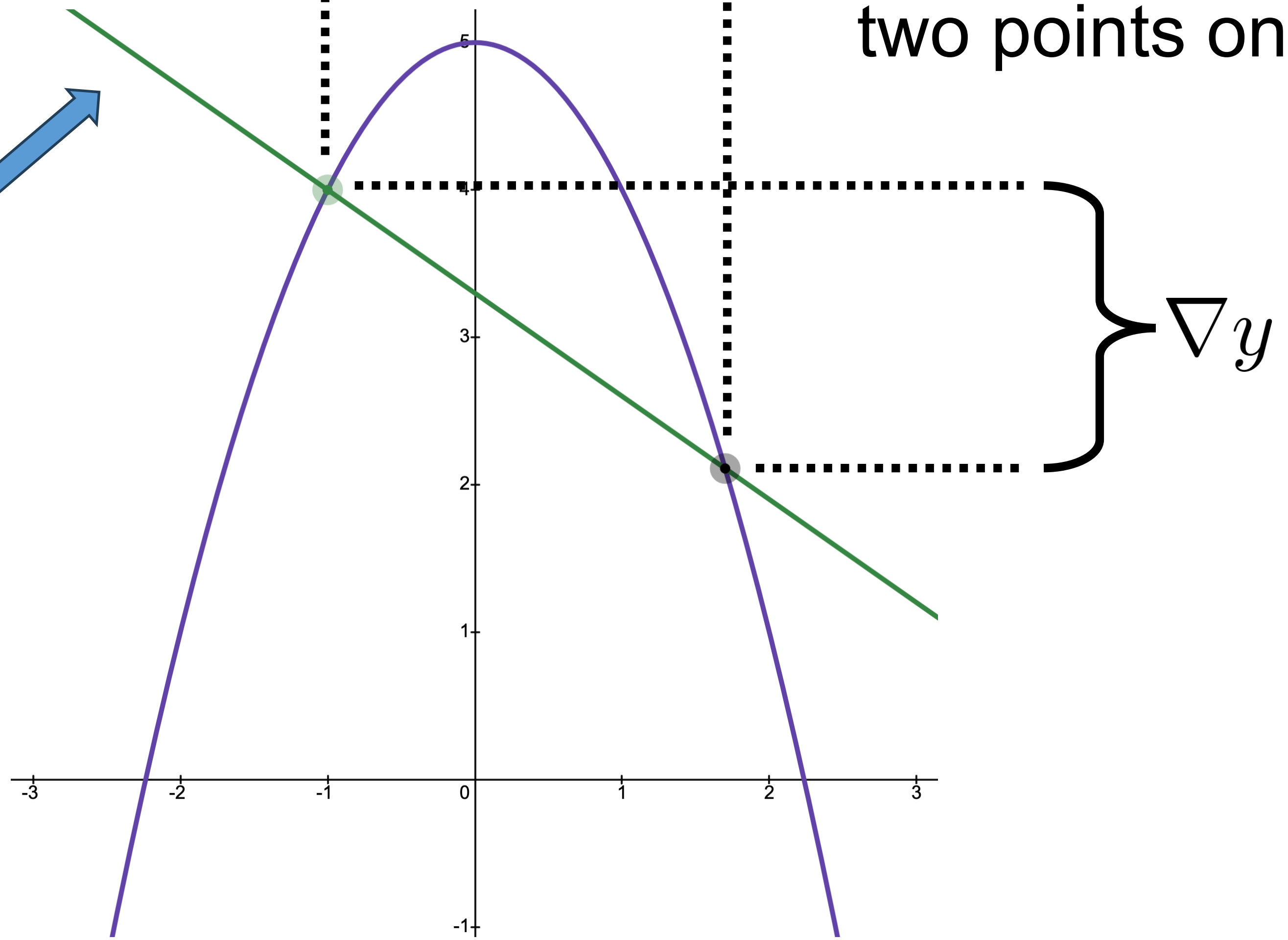
Slope



Secant

$$y = \frac{\nabla y}{\nabla x} x + b$$

Slope

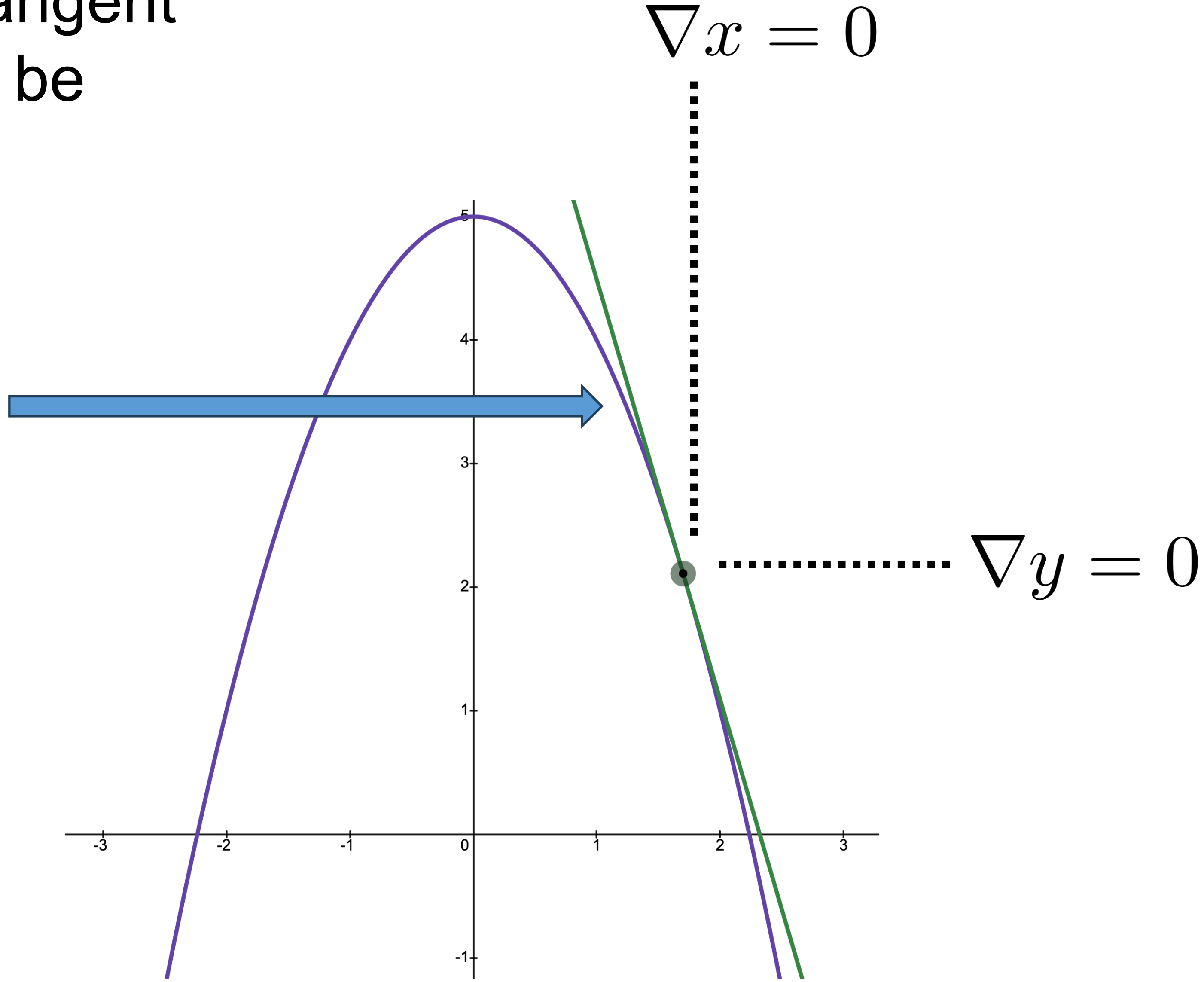


Think of the secant as the average slope between two points on a function.

Tangent

The slope of the tangent line appears to be undefined…

$$y = \frac{0}{0}x + b$$

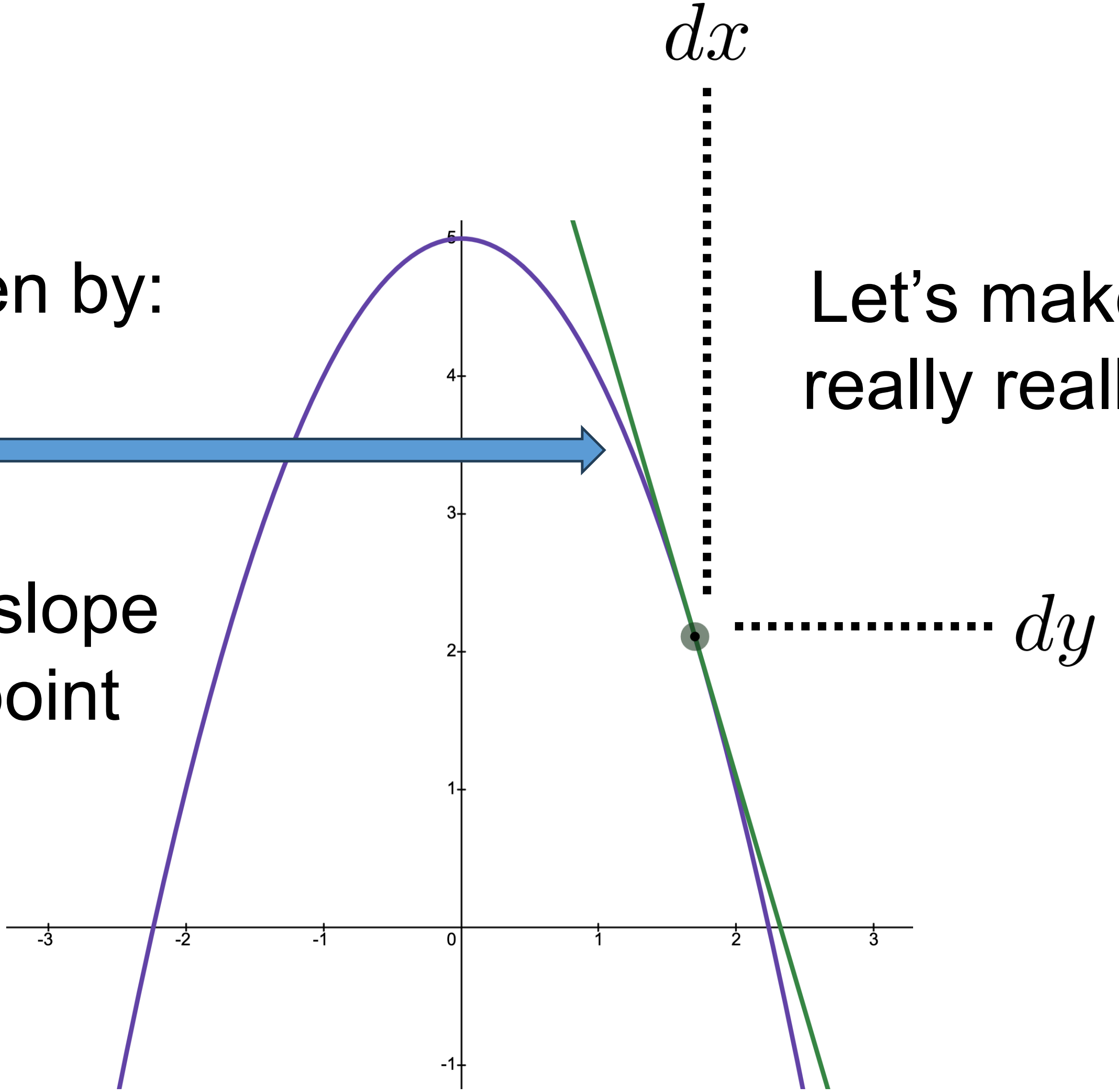


Derivative

The slope is now given by:

$$\frac{dy}{dx}$$

The *derivative* is the slope of the tangent at a point

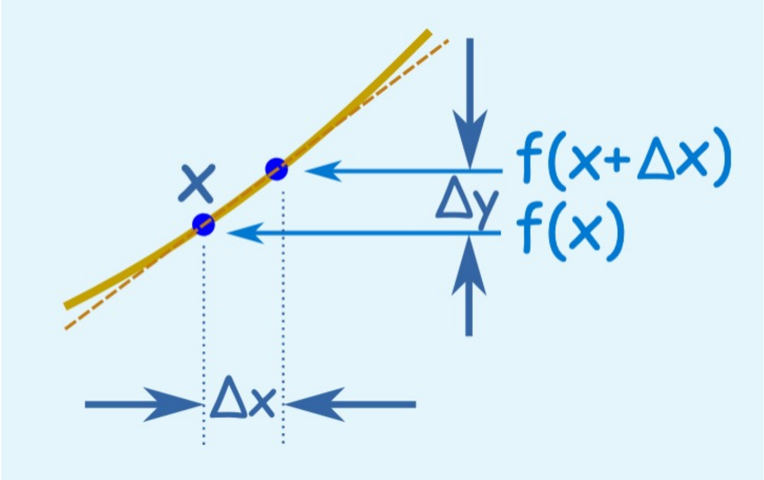


Let's make the difference really really really small…

How to Calculate a Derivative

To find the derivative of a function $y = f(x)$ we use the slope form:

$$\text{Slope} = \frac{\text{Change in } y}{\text{Change in } x} = \frac{\nabla y}{\nabla x}$$



X changes from: $x \to x + \nabla x$

Y changes from: $f(x) \to f(x + \nabla x)$

Start with the slope form: $\frac{\nabla y}{\nabla x} = \frac{f(x+\nabla x) - f(x)}{\nabla x}$

Simplify, and make ∇x shrink towards zero…

Example: Derivative

Example: the function $\mathbf{f(x) = x^2}$

The slope formula is:	$\frac{f(x+\Delta x) - f(x)}{\Delta x}$
Use $f(x) = x^2$:	$\frac{(x+\Delta x)^2 - x^2}{\Delta x}$
Expand $(x+\Delta x)^2$ to $x^2+2x\,\Delta x+(\Delta x)^2$:	$\frac{x^2 + 2x\,\Delta x + (\Delta x)^2 - x^2}{\Delta x}$
Simplify (x^2 and $-x^2$ cancel):	$\frac{2x\,\Delta x + (\Delta x)^2}{\Delta x}$
Simplify more (divide through by Δx):	$2x + \Delta x$
Then, **as Δx heads towards 0** we get:	$2x$

Result: the derivative of $\mathbf{x^2}$ is **2x**

In other words, the slope at x is **2x**

Source: mathsisfun.com

Derivative

Instead of saying “∇x heads towards zero” we write ”dx” so we have:

$$\frac{d}{dx}x^2 = 2x$$

But what does this actually mean?

- For function x^2 the slope (or rate of change) at any point is 2x
- So if x=2 then the slope is 2x=4
- If x=5 then the slope is 2x=10
- And so on…

Another Derivative Example

Example: What is $\frac{d}{dx}x^3$?

We know $\mathbf{f(x) = x^3}$, and can calculate $\mathbf{f(x+\Delta x)}$, so let's go:

The slope formula: $\frac{f(x+\Delta x) - f(x)}{\Delta x}$

Use $f(x) = x^3$: $\frac{(x+\Delta x)^3 - x^3}{\Delta x}$

Use $(x+\Delta x)^3 = x^3 + 3x^2\,\Delta x + 3x\,(\Delta x)^2 + (\Delta x)^3$

Replace $(x+\Delta x)^3$: $\frac{x^3 + 3x^2\,\Delta x + 3x\,(\Delta x)^2 + (\Delta x)^3 - x^3}{\Delta x}$

Simplify (x^3 and $-x^3$ cancel): $\frac{3x^2\,\Delta x + 3x\,(\Delta x)^2 + (\Delta x)^3}{\Delta x}$

Simplify more (divide through by Δx): $3x^2 + 3x\,\Delta x + (\Delta x)^2$

As Δx heads towards 0 we get: $3x^2$

Result: the derivative of $\mathbf{x^3}$ is $\mathbf{3x^2}$

Source: mathsisfun.com

Derivative Rules

In practice it is easier to memorize derivative rules…

Common Functions	Function	Derivative
Constant	c	0
Line	x	1
	ax	a
Square	x^2	$2x$
Square Root	$\sqrt{x}$	$(½)x^{-½}$
Exponential	e^x	e^x
	a^x	$\ln(a)\ a^x$
Logarithms	$\ln(x)$	$1/x$
	$\log_a(x)$	$1\ /\ (x \ln(a))$
Trigonometry (x is in radians)	$\sin(x)$	$\cos(x)$
	$\cos(x)$	$-\sin(x)$
	$\tan(x)$	$\sec^2(x)$
Inverse Trigonometry	$\sin^{-1}(x)$	$1/\sqrt{(1-x^2)}$
	$\cos^{-1}(x)$	$-1/\sqrt{(1-x^2)}$
	$\tan^{-1}(x)$	$1/(1+x^2)$

Rules	Function	Derivative
Multiplication by constant	cf	cf'
Power Rule	x^n	nx^{n-1}
Sum Rule	$f + g$	$f' + g'$
Difference Rule	$f - g$	$f' - g'$
Product Rule	fg	$f\,g' + f'\,g$
Quotient Rule	f/g	$\frac{f'\,g - g'\,f}{g^2}$
Reciprocal Rule	$1/f$	$-f'/f^2$
Chain Rule (using ′)	$f(g(x))$	$f'(g(x))g'(x)$
Chain Rule (as "Composition of Functions")	$f \circ g$	$(f' \circ g) \times g'$
Chain Rule (using $\frac{d}{dx}$)	$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$	

Source: mathsisfun.com

Derivative Rules: Power Rule

In practice it is easier to memorize derivative rules…

Common Functions	Function	Derivative
Constant	c	0
Line	x	1
	ax	a
Logarithms	ln(x)	1/x
	$\log_a(x)$	1 / (x ln(a))
Trigonometry (x is in radians)	sin(x)	cos(x)
	cos(x)	−sin(x)
	tan(x)	$\sec^2(x)$
Inverse Trigonometry	$\sin^{-1}(x)$	$1/\sqrt{(1-x^2)}$
	$\cos^{-1}(x)$	$-1/\sqrt{(1-x^2)}$
	$\tan^{-1}(x)$	$1/(1+x^2)$

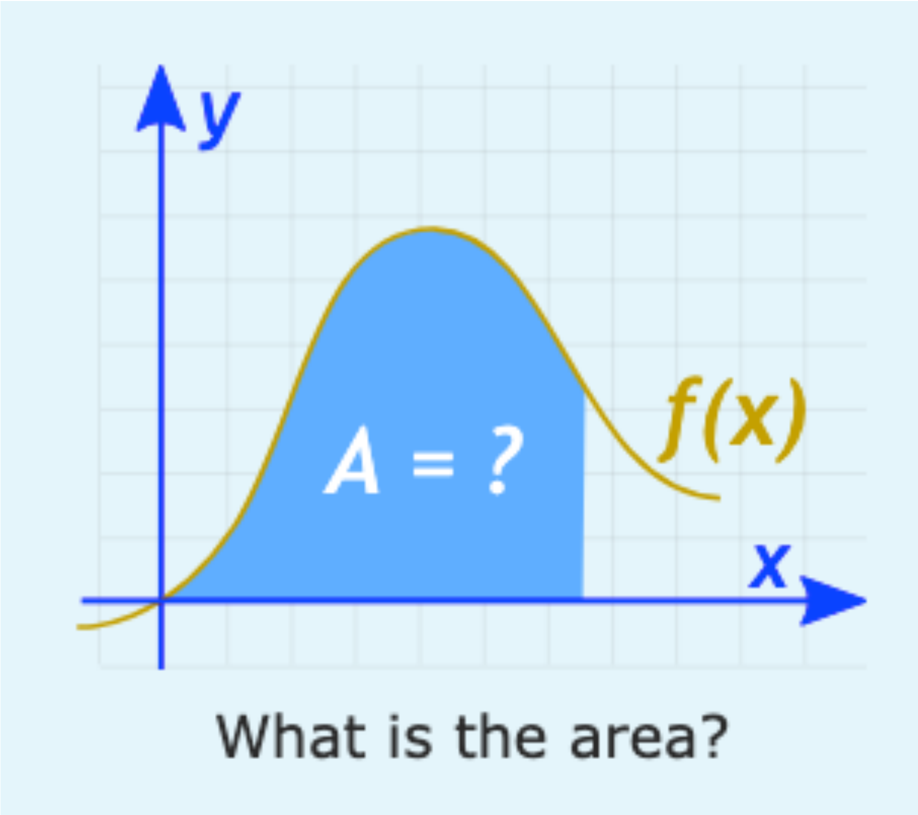
Rules	Function	Derivative
Multiplication by constant	cf	cf′
Power Rule	x^n	nx^{n-1}
Sum Rule	f + g	f′ + g′
Chain Rule (using ′)	f(g(x))	f′(g(x))g′(x)
Chain Rule (as "Composition of Functions")	$f \circ g$	$(f' \circ g) \times g'$
Chain Rule (using $\frac{d}{dx}$)	$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$	

Rules	Function	Derivative
Power Rule	x^n	nx^{n-1}

Source: mathsisfun.com

Integral

The <u>integral</u> computes the area under a function…

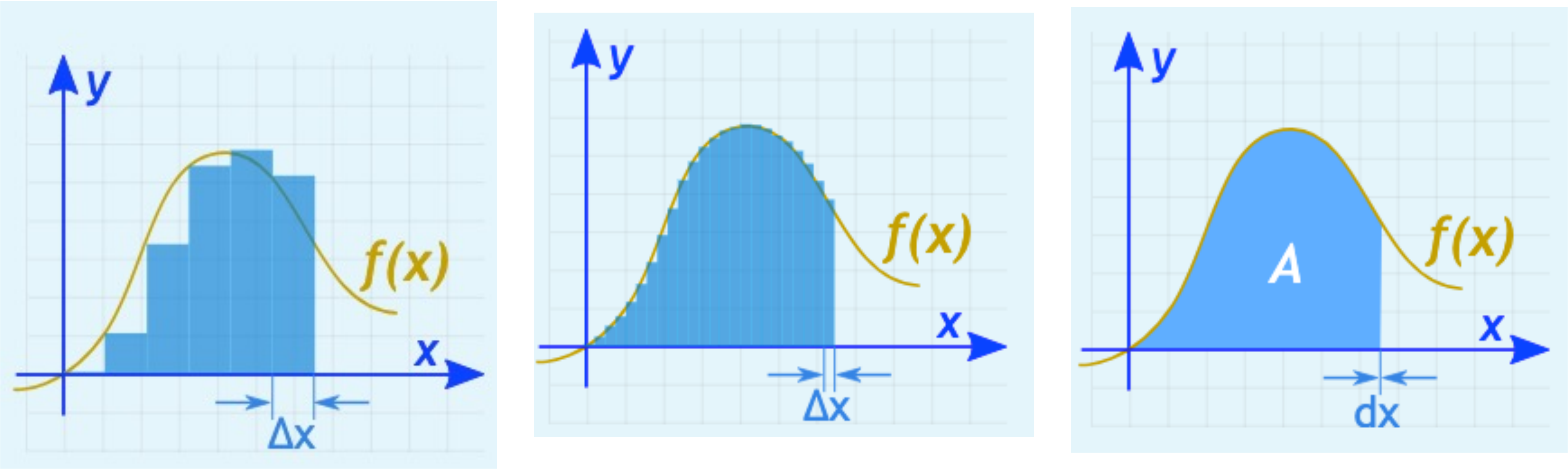


…it is also the reverse operation of a derivative.

Source: mathsisfun.com

Integration

We can view this as adding up smaller and smaller slices…



…where dx means slices are approaching zero width.

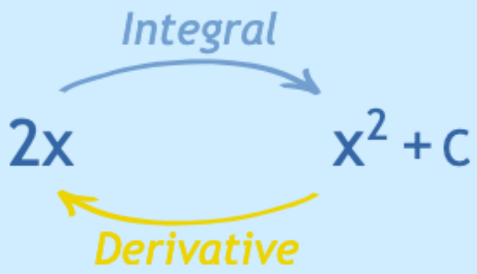
Source: mathsisfun.com

Integration

But there is a shortcut because the integral is the reverse of the derivative…

Example: 2x

An integral of 2x is x^2 ...



... because the derivative of x^2 is 2x

(More about "+C" later.)

Source: mathsisfun.com

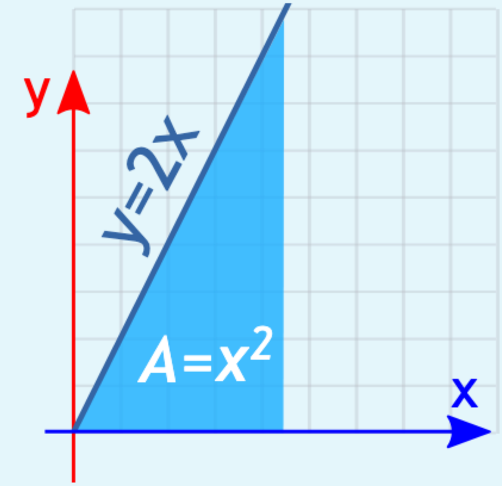
Integral Notation

Slices along x

$$\int 2x \, dx = x^2 + C$$

Integral Symbol

Function we want to integrate



Area of triangle $= \frac{1}{2}(\text{base})(\text{height}) = \frac{1}{2}(x)(2x) = x^2$

Source: mathsisfun.com

Plus C

$$\int 2x\,dx = x^2 + C$$

- C is “Constant of Integration”
- Captures all functions whose derivative is 2x
- Derivative of a constant C is 0 so…
  - Derivative of x^2 is 2x
  - Derivative of $x^2 + 4$ is 2x
  - Derivative of $x^2 + 99$ is 2x
  - Etc.
- Basically, just always add + C…

Rules of Integration

Common Functions	Function	Integral
Constant	$\int a\,dx$	$ax + C$
Variable	$\int x\,dx$	$x^2/2 + C$
Square	$\int x^2\,dx$	$x^3/3 + C$
Reciprocal	$\int (1/x)\,dx$	$\ln\lvert x\rvert + C$
Exponential	$\int e^x\,dx$	$e^x + C$
	$\int a^x\,dx$	$a^x/\ln(a) + C$
	$\int \ln(x)\,dx$	$x \ln(x) - x + C$
Trigonometry (x in radians)	$\int \cos(x)\,dx$	$\sin(x) + C$
	$\int \sin(x)\,dx$	$-\cos(x) + C$
	$\int \sec^2(x)\,dx$	$\tan(x) + C$
Rules	**Function**	**Integral**
Multiplication by constant	$\int cf(x)\,dx$	$c\int f(x)\,dx$
Power Rule ($n \neq -1$)	$\int x^n\,dx$	$\frac{x^{n+1}}{n+1} + C$
Sum Rule	$\int (f + g)\,dx$	$\int f\,dx + \int g\,dx$
Difference Rule	$\int (f - g)\,dx$	$\int f\,dx - \int g\,dx$
Integration by Parts	See Integration by Parts	
Substitution Rule	See Integration by Substitution	

Source: mathsisfun.com

Rules of Integration: Power Rule

Common Functions	Function	Integral
Constant	$\int a\,dx$	$ax + C$
Variable	$\int x\,dx$	$x^2/2 + C$
Square	$\int x^2\,dx$	$x^3/3 + C$
Reciprocal	$\int (1/x)\,dx$	$\ln\lvert x\rvert + C$
	$\int \sec^2(x)\,dx$	$\tan(x) + C$

Rules	Function	Derivative
Power Rule (n≠−1)	$\int x^n\,dx$	$\frac{x^{n+1}}{n+1} + C$

Rules	Function	Integral
Multiplication by constant	$\int cf(x)\,dx$	$c\int f(x)\,dx$
Power Rule (n≠−1)	$\int x^n\,dx$	$\frac{x^{n+1}}{n+1} + C$
Sum Rule	$\int (f + g)\,dx$	$\int f\,dx + \int g\,dx$
Difference Rule	$\int (f - g)\,dx$	$\int f\,dx - \int g\,dx$
Integration by Parts	See Integration by Parts	
Substitution Rule	See Integration by Substitution	

Source: mathsisfun.com

Example: Power Rule

Example: What is $\int x^3\, dx$?

The question is asking "what is the integral of x^3 ?"

We can use the Power Rule, where n=3:

$$\int x^n\, dx = \frac{x^{n+1}}{n+1} + C$$

$$\int x^3\, dx = \frac{x^4}{4} + C$$

Source: mathsisfun.com

Example: Power Rule

Example: What is $\int \sqrt{x}\, dx$?

$\sqrt{x}$ is also $\mathbf{x^{0.5}}$

We can use the Power Rule, where n=0.5:

$$\int x^n \, dx = \frac{x^{n+1}}{n+1} + C$$

$$\int x^{0.5} \, dx = \frac{x^{1.5}}{1.5} + C$$

Source: mathsisfun.com

Example: Multiplication by a Constant

Example: What is $\int 6x^2\,dx$?

We can move the 6 outside the integral:

$$\int 6x^2\,dx = 6\int x^2\,dx$$

And now use the Power Rule on x^2:

$$= 6\,\frac{x^3}{3} + C$$

Simplify:

$$= 2x^3 + C$$

Source: mathsisfun.com

Definite Integral

Same concept as a discrete summation…

$$\sum_{x=a}^{b} 2x$$

Discrete

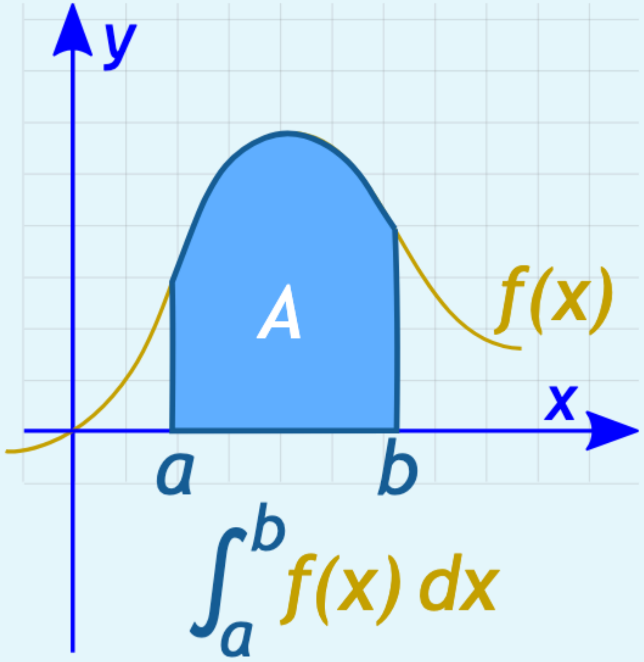
$$\int_a^b 2x \, dx$$

Continuous

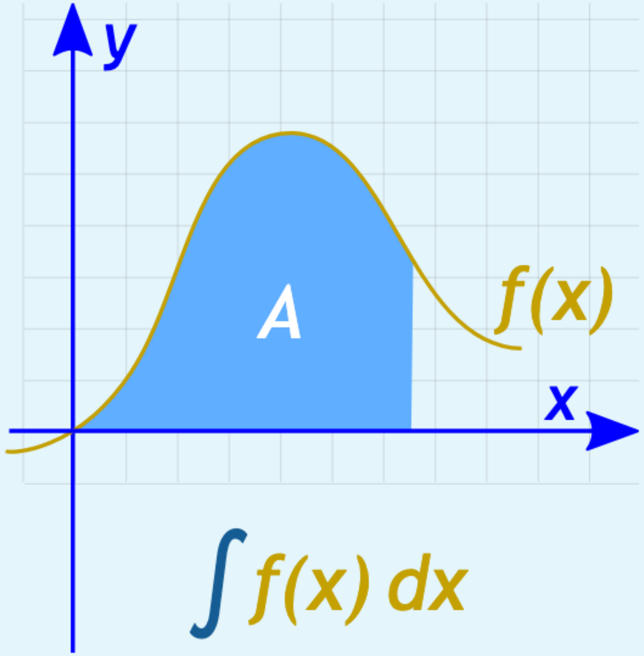
…integral symbol is a stylized "S" for "sum".

Definite Integral

***Definite Integral** can be used to find the area under a curve between two points…*



Definite Integral
(from **a** to **b**)



Indefinite Integral
(no specific values)

Source: mathsisfun.com

Definite Integral

First we compute the indefinite integral,

$$\int_a^b 2x\,dx = x^2 + C$$

Then we substitute and subtract the endpoints,

$$\left[x^2 + C\right]_a^b = b^2 - a^2$$

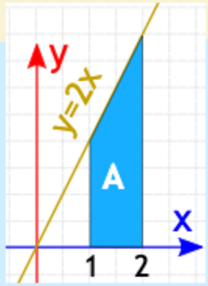
Notice that the constant of integration always cancels,

$$\left[x^2 + C\right]_a^b = (b^2 + C) - (a^2 + C) = b^2 - a^2$$

Example: Definite Integral

Example: What is

$$\int_1^2 2x\,dx$$



We are being asked for the **Definite Integral**, from 1 to 2, of **2x** dx

First we need to find the ***Indefinite* Integral**.

Using the Rules of Integration we find that $\int 2x\,dx = x^2 + C$

Now calculate that at 1, and 2:

- At x=1: $\int 2x\,dx = 1^2 + C$
- At x=2: $\int 2x\,dx = 2^2 + C$

Subtract:

$(2^2 + C) - (1^2 + C)$

$2^2 + C - 1^2 - C$

$4 - 1 + C - C = 3$

And "C" gets cancelled out ... so with **Definite Integrals we can ignore C**.

Source: mathsisfun.com

Example

What is the value of the definite integral:

$$\int_2^4 x^3 \, dx \ ?$$

A 56

B 60

C 68

D 240

Integral: Area Above / Below

The integral adds the area above the axis but subtracts the area below, for a "net value":

$$\int_a^b f(x)\,dx = (\text{Area above x axis}) - (\text{Area below x axis})$$

Source: mathsisfun.com

Integral: Adding Functions

The integral of **f+g** equals the integral of **f** plus the integral of **g**:

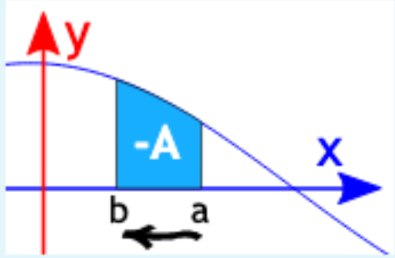
$$\int_a^b f(x) + g(x)\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx$$

Source: mathsisfun.com

Integral: Reversing the Interval

Reversing the direction of the interval gives the negative of the original direction.

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$$

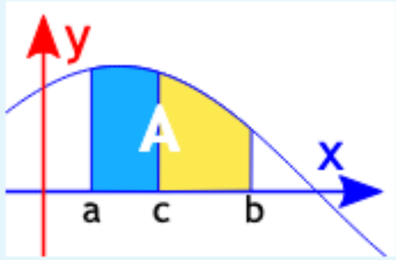


Source: mathsisfun.com

Integral: Adding Intervals

We can also add two adjacent intervals together:

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

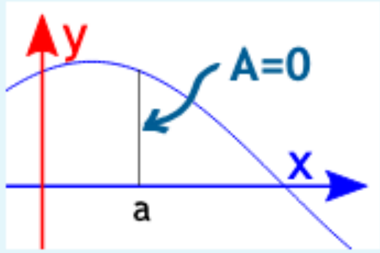


Source: mathsisfun.com

Integral: Interval of Zero Length

When the interval starts and ends at the same place, the result is zero:

$$\int_a^a f(x)\,dx = 0$$



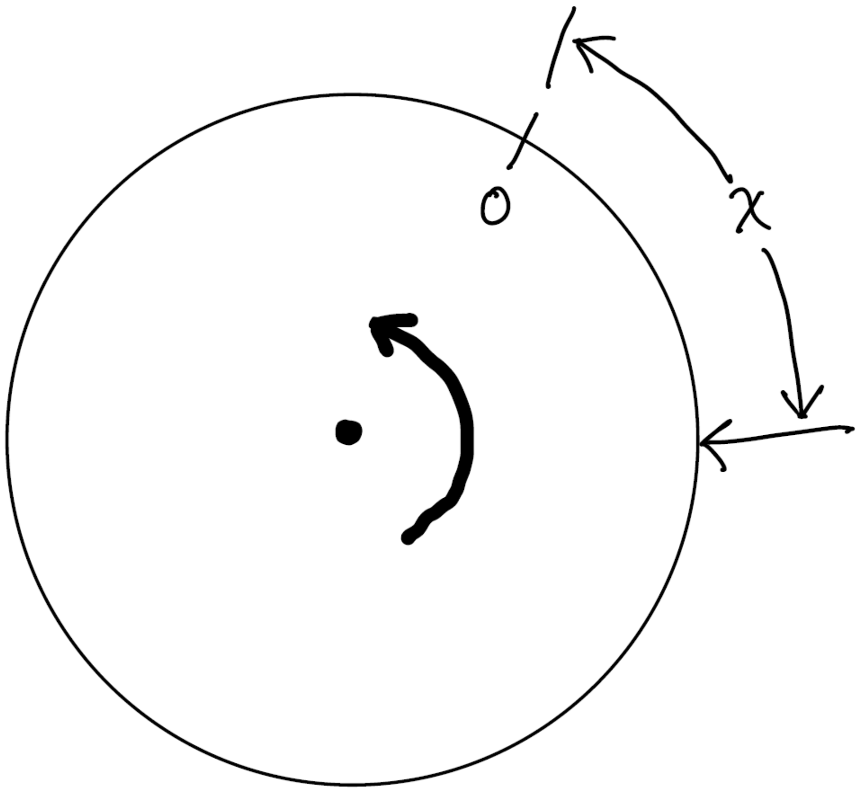
Source: mathsisfun.com

Outline

- Concepts of Calculus
- Continuous Probability Distributions
- Fundamental Rules of Probability

Continuous Probability

Experiment Spin continuous wheel and measure X displacement from 0



Question Assuming uniform probability, what is $P(X = x)$?

Continuous Probability

For continuous random variables, the probability of any outcome is zero:

$$P(X = x) = 0$$

Our preference is to define events as intervals, e.g. using the CDF:

$$P(X \leq x), \quad P(a < X \leq b), \quad P(X > c), \quad \text{etc.}$$

Note that when X is continuous,

$$P(a < X \leq b) = P(a < X < b) + P(X = b) = P(a < X < b)$$

thus, it does not matter if we include the endpoints.

Probability Density

Definition A function $p(X)$ is a **probability density function (PDF)** of a continuous random variable X if the following hold:

(a) It is nonnegative for all values in the support,

$$p(X = x) \geq 0$$

(b) The integral over all values in the support is 1,

$$\int p(X = x)\, dx = 1$$

(c) The probability of an interval is given by the integral of the PDF,

$$P(a < X < b) = \int_a^b p(X = x)\, dx$$

Continuous Probability

$$P(a \leq X < b) = \int_a^b p(X = x)\, dx$$

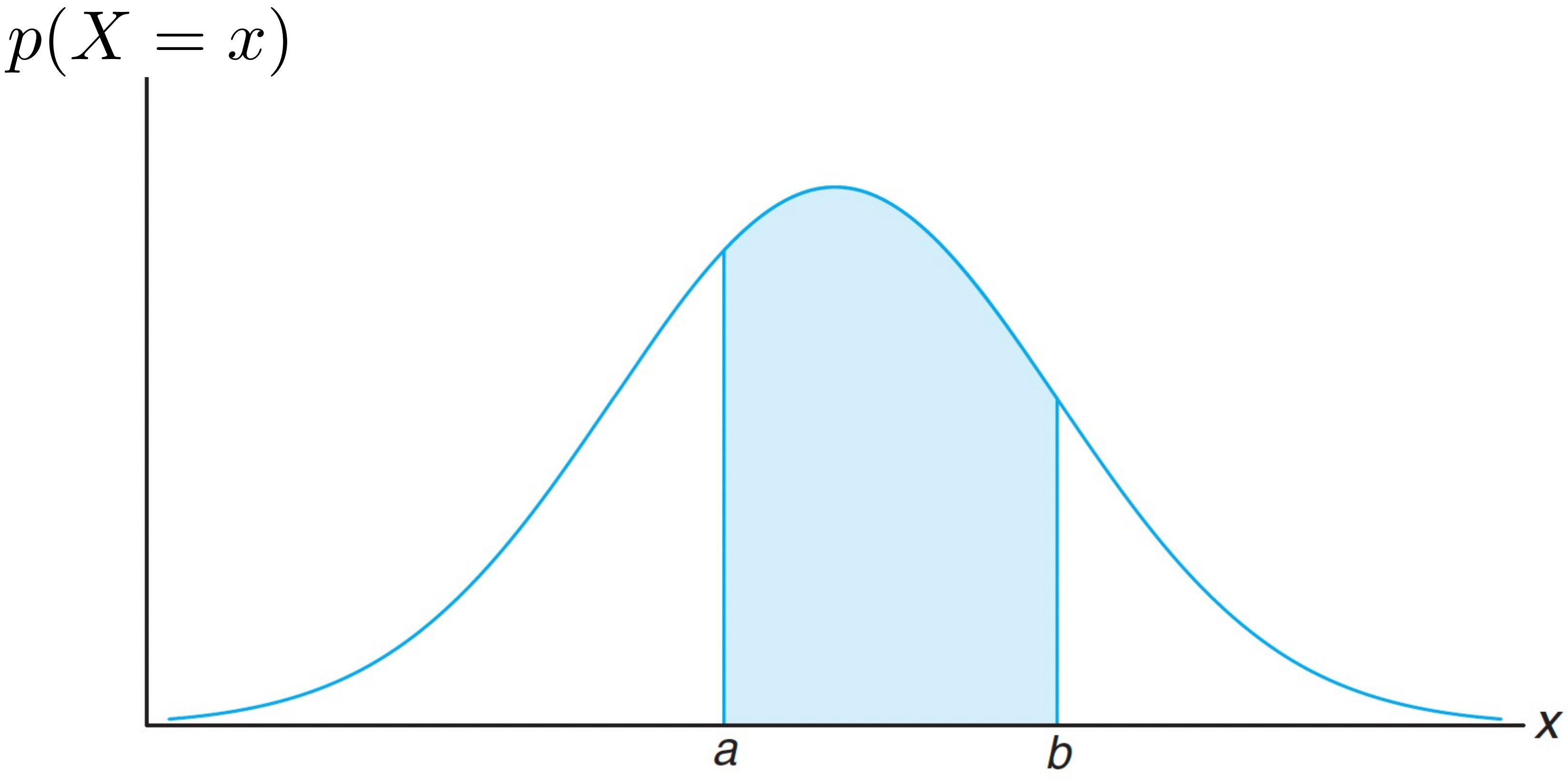


Figure 3.5: $P(a < X < b)$.

Continuous Probability Measures

- Can easily measure probability of closed intervals,

$$P(a \leq X < b) = P(X < b) - P(X < a)$$

- If X is *absolutely continuous* (i.e. differentiable) then,

Fundamental Theorem of Calculus

$$p(x) = \frac{dP(x)}{dx} \quad \text{and} \quad P(t) = \int_{-\infty}^{t} p(x)\, dx$$

Where $p(x)$ is the *probability density function* (PDF)

A word on notation…

The book uses slightly different notation from my slides…

The function $f(x)$ indicates the PMF or PDF:

$f(x) = P(X = x)$ for discrete X, and $f(x) = p(X = x)$ for continuous X.

The function $F(x)$ denotes the CDF for discrete and continuous RVs X:

$$F(x) = P(X < x)$$

…I will largely avoid the $f(x)$ and $F(x)$ notation except for some examples.

Outline

- Concepts of Calculus
- Continuous Probability Distributions
- Fundamental Rules of Probability

Continuous Probability

Most fundamental rules hold, replacing PMF with PDF/CDF…

Two RVs X & Y are **independent** if and only if,

$$p(x,y) = p(x)p(y) \qquad \text{or} \qquad P(X \le x, Y \le y) = P(X \le x)P(Y \le y)$$

Conditionally independent given Z iff,

Shorthand: $P(x) = P(X \le x)$

$$p(x,y \mid z) = p(x \mid z)p(y \mid z) \qquad \text{or} \qquad P(x,y \mid z) = P(x \mid z)P(y \mid z)$$

Probability chain rule,

$$p(x,y) = p(x)p(y \mid x) \qquad \text{and} \qquad P(x,y) = P(x)P(y \mid x)$$

Continuous Probability

…and by replacing summation with integration…

Law of Total Probability for continuous distributions,

$$p(x) = \int_{\mathcal{Y}} p(x, y)\, dy$$

Conditional density of a continuous random variable,

$$p(x \mid y) = \frac{p(x, y)}{p(y)}$$

Combining these we have the following:

$$p(x \mid y) = \frac{p(x, y)}{\int_{\mathcal{X}} p(x, y)\, dx}$$

Continuous Probability

Caution *Some technical subtleties arise in continuous spaces…*

For **discrete** RVs X & Y, the conditional

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

is **undefined** when P(Y=y) = 0 … no problem.

For **continuous** RVs we have,

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

but numerator and denominator are 0/0.

…we will just work with the conditional PDF for now…

Fundamental Rules of Probability

Law of total probability

$$p(Y) = \int p(Y, X = x)\, dx$$

- p(y) is a **marginal** distribution
- This is called **marginalization**

Proof

$$\int p(Y, X = x)\, dx = \int p(Y) p(X = x \mid Y)\, dx \quad \text{(chain rule)}$$

$$= p(Y) \int p(X = x \mid Y)\, dx \quad \text{(distributive property)}$$

$$= p(Y) \quad \text{(PDF Integrates to 1)}$$

Fundamental Rules of Probability

Given two continuous RVs X and Y the **conditional density** is:

$$p(X \mid Y) = \frac{p(X,Y)}{p(Y)}$$

By the law of total probability, we also have the definition:

$$p(X \mid Y) = \frac{p(X,Y)}{\int p(X{=}x,Y)\,dx}$$

Independence of RVs

Definition *Two random variables* X *and* Y *are* *independent* *if and only if,*

$$p(X = x, Y = y) = p(X = x)p(Y = y)$$

- This must hold for all values x and y.
- If for any values x and y,

$$p(X = x, Y = y) \neq p(X = x)p(Y = y)$$

  then X and Y are **dependent.**
- Example: Rolling two dice, each die is independent of the other
- Independence is *symmetric:* if X is independent of Y then Y is independent of X
- Equivalent definition of independence: $p(X \mid Y) = p(X)$

Independence of RVs

Definition RVs $X_1, X_2, \ldots, X_N$ are mutually independent if and only if,

$$p(X_1 = x_1, \ldots, X_N = x_N) = \prod_{i=1}^{N} p(X_i = x_i)$$

In words: If a set of random variables is independent, then their joint probability is a product of their marginals.

Independence of RVs

Definition *Two random variables* X *and* Y *are* *<u>conditionally independent</u>* *given* Z *if and only if,*

$$p(X = x, Y = y \mid Z = z) = p(X = x \mid Z = z)p(Y = y \mid Z = z)$$

for all values x, y, and z.

- N RVs conditionally independent, given Z, if and only if:

$$p(X_1 = x_1, \ldots, X_N = X_N \mid Z = z) = \prod_{i=1}^{N} p(X_i = x_i \mid Z = z)$$

- Equivalent def'n of conditional independence: $p(X \mid Y, Z) = p(X \mid Z)$
- Conditional independence is symmetric