Statistical Theory and Modeling (ST2601) Integration

Mattias Villani

Department of Statistics
Stockholm University

mattiasvillani.com @matvil @matvil mattiasvillani

Overview

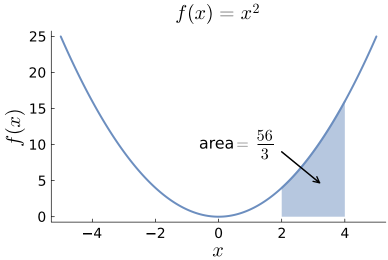
- **Integration**

Integration

- We often need to compute **the area under a function**.
- Statistics: **probabilities for continuous random variables**

$$\Pr(a \leq X \leq b)$$

are areas under the probability density function $f(x)$.



Rectangle sum to approximate areas

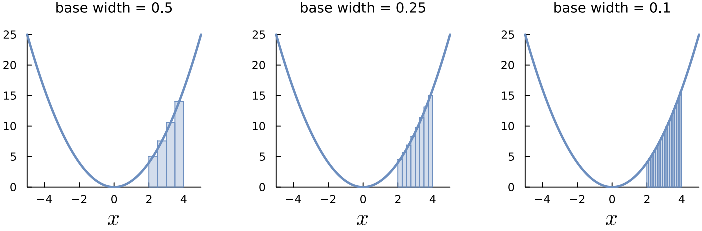
- Approximate area under $f(x)$ over $[a, b]$ by a **rectangle sum**

$$\sum_{i=1}^{n} f(x_i^\star)\Delta x_i$$

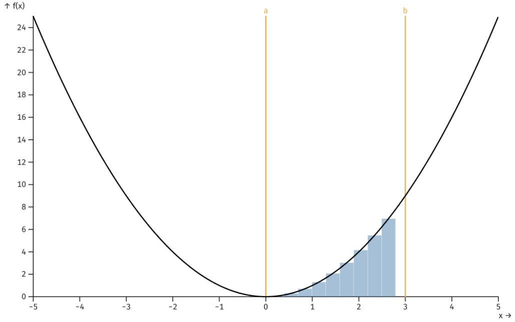
- **Partitioning** of the interval $[a, b]$

$$x_0 = a < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$$

- $x_i^\star$ is some value in the ith bin, for example the midpoint.

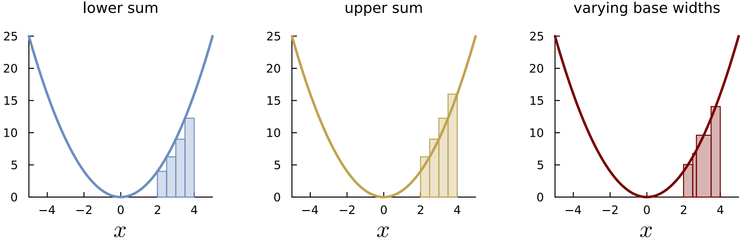


Integration widget



Many choices for rectangle design

- Width of rectangles Δx_i
- Height of rectangles (midpoint, lower or upper sum)
- Equal width?



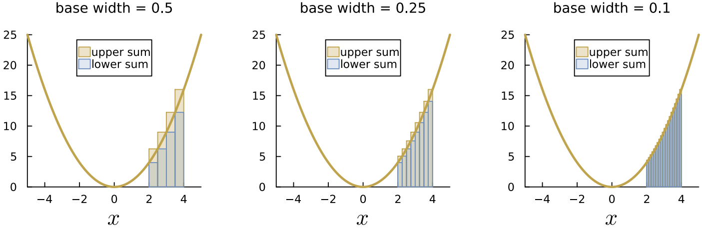
The Riemann integral

- Idea to compute area under $f(x)$ over $[a, b]$
  - Approximate by *both* lower and upper sum of rectangles.
  - Let the rectangle widths approach zero $\Delta x_i \to 0$
  - If **lower and upper sum converge** to the same value, then the function is Riemann **integrable** with integral $\int_a^b f(x)\mathrm{d}x$

$$\sum_{i=1}^{n} f(x_i^\star)\Delta x_i \to \int_a^b f(x)\mathrm{d}x$$

- The notation is really thoughtful:
  - The integral sign $\int$ looks like the letter *s* as in *sum*.
  - The $\mathrm{d}x$ is a small version of Δx (Δ is capital D in greek).

Lower and upper sums converging



The fundamental theorem of calculus

- Computing integrals by limiting rectangle sums is messy.
- The **anti-derivative** is a life-saver

Definition. *A function $F(x)$ is the* ***anti-derivative*** *to the function $f(x)$ if*

$$F'(x) = f(x), \textit{ for all } x$$

- The second **fundamental theorem of calculus**

Theorem 1. *If $f(x)$ is integrable on $[a, b]$ and $F(x)$ is an anti-derivative of $f(x)$, then*

$$\int_a^b f(x)\mathrm{d}x = F(b) - F(a),$$

Definite vs indefinite integrals

- The anti-derivative is also called an **indefinite integral**

$$F(x) = \int f(x)\mathrm{d}x$$

- A **definite integral** is the integral over a given interval $[a, b]$

$$\int_a^b f(x)\mathrm{d}x$$

- A definite integral is a *number*.
- An indefinite integral (anti-derivative) is a *function.*

Improper/generalized integrals

- Two general cases:
  1. The function $f(x)$ is **unbounded** for some x.
  2. One or both the integral endpoints a and b is $\pm\infty$

$$\int_{-\infty}^{b} f(x)\mathrm{d}x \qquad \int_{a}^{\infty} f(x)\mathrm{d}x \qquad \int_{-\infty}^{\infty} f(x)\mathrm{d}x$$

- Example of 1:

$$\int_0^1 \frac{1}{\sqrt{x}}\mathrm{d}x$$

- Example of 2: density function of Exponential distribution

$$\int_0^\infty \lambda e^{-\lambda x}\mathrm{d}x$$

- The types of improper integrals are handled as **limits**:

$$\int_a^\infty f(x)\mathrm{d}x = \lim_{b\to\infty}\int_a^b f(x)\mathrm{d}x$$

Diverging/converging improper integrals

- An improper integral can **diverge**

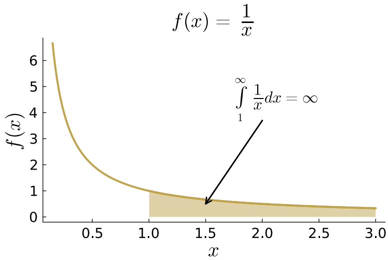
$$\int_1^\infty \frac{1}{x}\,\mathrm{d}x = \infty$$

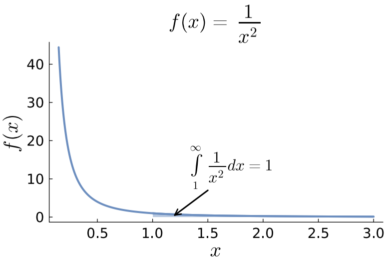
$f(x) = \frac{1}{x}$ does not go fast enough to zero as $x \to \infty$.

- Or **converge**

$$\int_1^\infty \frac{1}{x^2}\,\mathrm{d}x = 1$$

$f(x) = \frac{1}{x^2}$ goes to zero fast enough.





Integrals for common functions

Anti-derivatives of elementary functions

$f(x)$	$F(x)$	comment
x^n	$\frac{1}{n+1}x^{n+1}$	for $n \neq -1$
e^{ax}	$\frac{1}{a}e^{ax}$	for $a \neq 0$
$\frac{1}{x}$	$\ln\lvert x\rvert$	
a^x	$\frac{a^x}{\ln a}$	
$\sin x$	$-\cos x$	
$\cos x$	$\sin x$	

Integral rules for combined functions

Integrals for combinations of functions

Constant rule $\int_a^b kf(x)\,\mathrm{d}x = k\int_a^b f(x)\,\mathrm{d}x$ for constant k

Sum rule $\int_a^b (f(x)+g(x))\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x$

Product rule $\int_a^b f(x)g'(x)\,\mathrm{d}x = \left[f(x)g(x)\right]_a^b - \int_a^b f'(x)g(x)\,\mathrm{d}x$