Calculus

Last modified: 2026-09-28 23:45:45 (PDT)

Derivatives

Theorem 0.1 (Constant rule). *If c is constant with respect to x, then*

$$\frac{\partial}{\partial x} c = 0$$

Theorem 0.2 (Constant multiple rule). *If a is constant with respect to x and y is a differentiable function of x, then:*

$$\frac{\partial}{\partial x} ay = a\frac{\partial y}{\partial x}$$

Theorem 0.3 (Power rule). *For every real number q and every $x > 0$:*

$$\frac{\partial}{\partial x} x^q = qx^{q-1}$$

When q is a positive integer, the same formula holds for every real x.

Theorem 0.4 (Derivative of natural logarithm). *For every $x > 0$:*

$$\log'\{x\} = \frac{1}{x} = x^{-1}$$

Theorem 0.5 (derivative of exponential). *For every real x:*

$$\exp'\{x\} = \exp\{x\}$$

Theorem 0.6 (Product rule)**.** *If a and b are differentiable functions of x, then*

$$(ab)' = ab' + ba'$$

Theorem 0.7 (Quotient rule)**.** *If a and b are differentiable functions of x and $b \neq 0$, then*

$$(a/b)' = a'/b - (a/b^2)b'$$

Theorem 0.8 (Chain rule)**.** *If a is a differentiable function of b, and b is a differentiable function of c, then a is a differentiable function of c, and*

$$\begin{aligned} \frac{da}{dc} &= \frac{da}{db}\frac{db}{dc} \\ &= \frac{db}{dc}\frac{da}{db} \end{aligned}$$

or in Lagrange's notation, if g is differentiable at x and f is differentiable at $g(x)$:

$$(f(g(x)))' = g'(x)f'(g(x))$$

Corollary 0.1 (Chain rule for logarithms)**.** *If f is differentiable at x and $f(x) > 0$, then*

$$\frac{d}{dx}\log\{f(x)\} = \frac{f'(x)}{f(x)}$$

Proof

Proof. Apply Theorem 0.8 and Theorem 0.4:

$$\begin{aligned} \frac{d}{dx}\log\{f(x)\} &= f'(x) \cdot \log'\{f(x)\} && \text{(chain rule, with } g = f \text{ and outer function } \log) \\ &= f'(x) \cdot \frac{1}{f(x)} && \text{(derivative of } \log\text{, valid because } f(x) > 0) \\ &= \frac{f'(x)}{f(x)} && \text{(multiply)} \end{aligned}$$

□

Integration

Integration is the inverse operation of differentiation: it recovers a function from its derivative and accumulates quantities such as areas, totals, and probabilities. We begin with antiderivatives, then state basic integration rules, and conclude with the Fundamental Theorem of Calculus and a worked example from probability.

Antiderivatives

Definition 0.1 (Antiderivative)**.** A function F is an **antiderivative** of f on an interval I if:

$$\frac{\partial}{\partial x} F(x) = f(x), \quad \forall x \in I$$

(Larson and Edwards 2018, sec. 4.1, pp. 248–249)

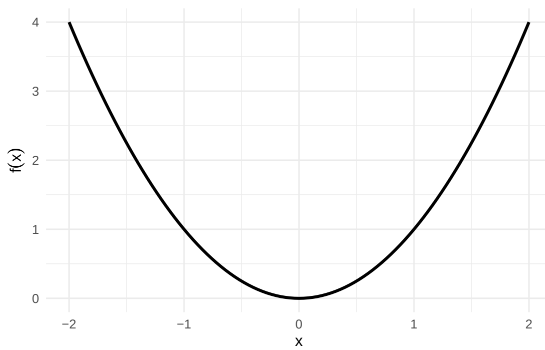
Definition 0.2 (Indefinite integral)**.** The **indefinite integral** of f is the family of all antiderivatives (Definition 0.1) of f:

$$\int f(x)\, dx = F(x) + C$$

where F is any one antiderivative of f and C is an arbitrary constant of integration. (Larson and Edwards 2018, sec. 4.1, pp. 248–249)

Example 0.1 (Antiderivative of a power function)**.** For $f(x) = x^2$, an antiderivative is $F(x) = \frac{x^3}{3}$, since $\frac{\partial}{\partial x} \frac{x^3}{3} = x^2 = f(x)$.
Adding any constant C gives another antiderivative; for example, with $C = 7$, $F(x) = \frac{x^3}{3} + 7$ also satisfies $F'(x) = x^2$, since adding a constant does not change the derivative. See Figure 1.

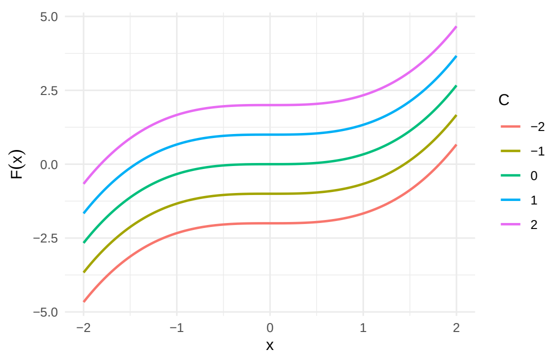
```r
ggplot2::ggplot() +
  ggplot2::geom_function(fun = \(x) :
  ggplot2::labs(x = "x", y = express
  ggplot2::theme_minimal()
```



(a) The function $f(x) = x^2$.

```r
x_seq <- seq(x_lim[1], x_lim[2], length.out = 200)
df <- do.call(rbind, lapply(C_vals, \(C) {
  data.frame(
    x = x_seq,
    y = x_seq^3 / 3 + C,
    C = factor(C)
  )
}))

ggplot2::ggplot(df, ggplot2::aes(x = x, y = y, color =
  ggplot2::geom_line(linewidth = 0.8) +
  ggplot2::labs(x = "x", y = expression(F(x)), color =
  ggplot2::theme_minimal()
```



(b) Family of antiderivatives $F(x) = x^3/3 + C$.

Figure 1: The function $f(x) = x^2$ and five antiderivatives $F(x) = x^3/3 + C$ for $C \in \{-2, -1, 0, 1, 2\}$. Each antiderivative has the same derivative f; they differ only by a vertical shift.

Theorem 0.9 (Basic integration rules)**.** *Each antiderivative in the table is defined only up to an arbitrary constant C (see Definition 0.1); the table omits $+C$ from every row for brevity.*

Function $f(x)$	*Antiderivative $F(x)$*	*Condition*
c	cx	—
x^n	$\dfrac{x^{n+1}}{n+1}$	$n \neq -1$
$\dfrac{1}{x}$	$\log\{\lvert x\rvert\}$	$x \neq 0$
e^x	e^x	*(self-antiderivative)*

e^{cx}	$\frac{1}{c}e^{cx}$	$c \neq 0$
$c \cdot f(x)$	$c \cdot F(x)$	—
$f(x) + g(x)$	$F(x) + G(x)$	—

The first two rows and the bottom two rows (linearity) are from (Larson and Edwards 2018, sec. 4.1, p. 250 "Basic Integration Rules"); $1/x$ is from (Larson and Edwards 2018, sec. 5.2, Theorem 5.5, p. 324); e^x and e^{cx} are from (Larson and Edwards 2018, sec. 5.4, Theorem 5.12, p. 346).

Example 0.2 (Antiderivative of $3x^2 - 1$)**.** By the power rule ($n = 2$) and linearity from Theorem 0.9:

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

Verify by differentiating: $\frac{\partial}{\partial x}(x^3 - x + C) = 3x^2 - 1 = f(x)$, as required.

Regularity Conditions

Definition 0.3 (Differentiable function)**.** A function f is **differentiable at** $x = c$ if the limit

$$f'(c) \stackrel{\text{def}}{=} \lim_{h \to 0} \frac{f(c+h) - f(c)}{h}$$

exists and is finite.
(Larson and Edwards 2018, sec. 2.1, p. 100)

Definition 0.4 (Differentiable on an interval)**.** A function f is **differentiable on** an interval if it is differentiable (Definition 0.3) at every interior point of the interval; at a closed endpoint, the appropriate one-sided derivative is used.
(Larson and Edwards 2018, sec. 2.1, p. 100)

Definition 0.5 (Continuous function)**.** A function f is **continuous at** $x = c$ if all three conditions hold:

1. $f(c)$ is defined,
2. $\lim_{x \to c} f(x)$ exists, and
3. $\lim_{x \to c} f(x) = f(c)$.

(Larson and Edwards 2018, sec. 1.4, p. 73)

Definition 0.6 (Continuous on a closed interval)**.** A function f is **continuous on** a closed interval $[a, b]$ if all three conditions hold:

1. f is continuous (Definition 0.5) at every point of the open interval (a, b),
2. $\lim_{x \to a^+} f(x) = f(a)$, and
3. $\lim_{x \to b^-} f(x) = f(b)$.

(Larson and Edwards 2018, sec. 1.4, p. 73)

Example 0.3 (Continuity on $[0, 1]$ uses one-sided limits at the endpoints)**.** Let $f(x) = \sqrt{x}$, defined for $x \geq 0$. Because f is undefined for $x < 0$, only the right-hand limit of f at 0 makes sense, and Definition 0.6 asks only for that one-sided limit at the endpoint 0. Here f is continuous at every point of $(0, 1)$, $\lim_{x \to 0^+} \sqrt{x} = 0 = f(0)$, and $\lim_{x \to 1^-} \sqrt{x} = 1 = f(1)$, so f is continuous on $[0, 1]$ (Definition 0.6).

Definition 0.7 (Partition of an interval)**.** A **partition** $\mathcal{P}$ of a closed interval $[a, b]$ is a finite list of points

$$a = x_0 < x_1 < \cdots < x_n = b.$$

It splits $[a, b]$ into the n subintervals $[x_{i-1}, x_i]$, of widths $\Delta x_i \stackrel{\text{def}}{=} x_i - x_{i-1}$. (Larson and Edwards 2018, sec. 4.3)

Example 0.4 (A partition of $[0, 1]$)**.** The points $0 < 0.25 < 0.5 < 1$ form a partition of $[0, 1]$ with $n = 3$ subintervals, of widths $\Delta x_1 = 0.25$, $\Delta x_2 = 0.25$, and $\Delta x_3 = 0.5$.

Definition 0.8 (Mesh of a partition)**.** The **mesh** of a partition $\mathcal{P}$ (Definition 0.7) is its largest subinterval width,

$$\|\mathcal{P}\| \stackrel{\text{def}}{=} \max_i \Delta x_i.$$

(Larson and Edwards 2018, sec. 4.3)

Example 0.5 (The mesh of a partition of $[0, 1]$)**.** For the partition of $[0, 1]$ in Example 0.4, with widths $\Delta x_1 = 0.25$, $\Delta x_2 = 0.25$, and $\Delta x_3 = 0.5$, the mesh is the largest of these widths, $\|\mathcal{P}\| = 0.5$.

Definition 0.9 (Riemann integral)**.** Let f be a bounded function on $[a, b]$. For each partition $\mathcal{P}$ of $[a, b]$ (Definition 0.7), choose a sample point x_i^* in each subinterval $[x_{i-1}, x_i]$. The **Riemann integral** of f over $[a, b]$ is the limit as the mesh (Definition 0.8) shrinks to zero:

$$\int_a^b f(x)\,dx \stackrel{\text{def}}{=} \lim_{\|\mathcal{P}\|\to 0} \sum_{i=1}^{n} f(x_i^*)\,\Delta x_i,$$

when that limit exists and has the same value for every choice of the partitions and of the sample points x_i^*.
(Larson and Edwards 2018, sec. 4.3, p. 272)

Definition 0.10 (Riemann integrable)**.** A bounded function f is **Riemann integrable on** $[a, b]$ if its Riemann integral (Definition 0.9) over $[a, b]$ exists and is finite.
(Larson and Edwards 2018, sec. 4.3, p. 272)

Definition 0.11 (Equal-width Riemann sum)**.** For a bounded function f on $[a, b]$ and a positive integer n, split $[a, b]$ into n subintervals of equal width $\Delta x \stackrel{\text{def}}{=} (b - a)/n$, and let x_i^* be any point in the i-th subinterval. The **equal-width Riemann sum** is

$$S_n \stackrel{\text{def}}{=} \sum_{i=1}^{n} f(x_i^*)\,\Delta x.$$

Example 0.6 (An equal-width Riemann sum)**.** Let $f(x) = x^2$ on $[0, 1]$, with $n = 2$, so $\Delta x = 1/2$, and take each sample point at the right end of its subinterval: $x_1^* = \frac{1}{2}$ and $x_2^* = 1$. Then

$$\begin{aligned} S_2 &= f\left(\tfrac{1}{2}\right)\cdot\tfrac{1}{2} + f(1)\cdot\tfrac{1}{2} && \text{(equal-width Riemann sum with } n = 2) \\ &= \tfrac{1}{4}\cdot\tfrac{1}{2} + 1\cdot\tfrac{1}{2} && \text{(evaluate } f(x) = x^2) \\ &= \tfrac{5}{8} && \text{(add)} \end{aligned}$$

Before stating the Fundamental Theorem of Calculus, we record two prerequisite results. The usual statement of the Fundamental Theorem of Calculus assumes that the integrand f is continuous on $[a, b]$; continuity is sufficient there, though not necessary. The two results are "differentiability implies continuity", which says where continuity comes from, and "continuity implies integrability", which says what continuity buys us.

Theorem 0.10 (Differentiability implies continuity)**.** *If f is differentiable at $x = c$, then f is continuous at $x = c$.*

(Larson and Edwards 2018, Theorem 2.1, p. 106)

Proof

Proof. Because $f'(c)$ exists, $f(c)$ is defined, and:

$$\begin{aligned}
\lim_{h\to 0}(f(c+h)-f(c)) &= \lim_{h\to 0}\left(\frac{f(c+h)-f(c)}{h}\cdot h\right) && \text{(multiply and divide by } h\neq 0\text{)}\\
&= \left(\lim_{h\to 0}\frac{f(c+h)-f(c)}{h}\right)\cdot\left(\lim_{h\to 0} h\right) && \text{(limit of a product, both limits exist)}\\
&= f'(c)\cdot 0 && \text{(definition of } f'(c)\text{)}\\
&= 0 && \text{(multiply)}
\end{aligned}$$

So $\lim_{h\to 0} f(c+h) = f(c)$, which is $\lim_{x\to c} f(x) = f(c)$ with $x = c+h$; all three conditions of Definition 0.5 hold. □

Example 0.7 (Differentiable, hence continuous: $x^3 - x$)**.** $f(x) = x^3 - x$ is differentiable everywhere (with derivative $f'(x) = 3x^2 - 1$), so by Theorem 0.10 it is continuous everywhere.

Example 0.8 (Continuous but not differentiable: $|x|$)**.** The absolute-value function $f(x) = |x|$ is continuous at $x = 0$ ($\lim_{x\to 0}|x| = 0 = |0|$), but it is not differentiable at $x = 0$: the left-derivative is -1 and the right-derivative is $+1$.
This counterexample shows that the converse of Theorem 0.10 fails: continuity does not imply differentiability. See Figure 2.

```r
ggplot2::ggplot() +
  ggplot2::geom_function(fun = abs, xlim = c(-2, 2), linewidth = 1) +
  ggplot2::geom_point(ggplot2::aes(x = 0, y = 0), size = 3) +
  ggplot2::labs(x = "x", y = expression(f(x) == abs(x))) +
  ggplot2::theme_minimal()
```

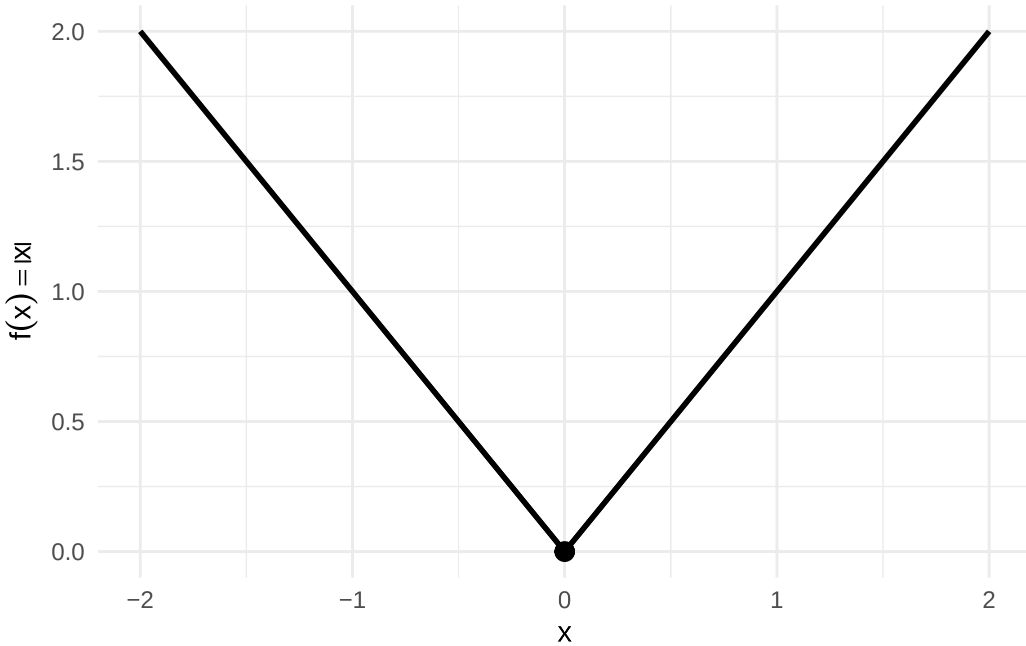


Figure 2: $f(x) = |x|$ has a sharp corner at $x = 0$ (not differentiable there) but is continuous everywhere: no gaps or jumps.

Theorem 0.11 (Continuity implies integrability)**.** *If f is continuous on the closed interval $[a, b]$, then f is integrable on $[a, b]$ (i.e., the Riemann integral $\int_a^b f(x)\,dx$ exists and is finite).*
(Larson and Edwards 2018, Theorem 4.4, p. 272)

Example 0.9 (Continuous, hence integrable: polynomials)**.** Every polynomial is continuous on $\mathbb{R}$, so by Theorem 0.11 every polynomial is integrable on every closed interval $[a, b]$.

Example 0.10 (Integrable but not continuous: a step function)**.** Let $f(x) = 0$ for $x < \frac{1}{2}$ and $f(x) = 1$ for $x \geq \frac{1}{2}$. Then f is discontinuous at $x = \frac{1}{2}$, but it is integrable on $[0, 1]$:

$$\int_0^1 f(x)\,dx = \int_0^{1/2} 0\,dx + \int_{1/2}^1 1\,dx = 0 + \tfrac{1}{2} = \tfrac{1}{2}.$$

This counterexample shows that the converse of Theorem 0.11 fails: integrability does not imply continuity. See Figure 3.

```r
step_df <- data.frame(
  x = c(0, 0.5, 0.5, 1),
  y = c(0, 0, 1, 1),
  segment = c("left", "left", "right", "right")
)

ggplot2::ggplot() +
  ggplot2::geom_rect(
    ggplot2::aes(xmin = 0.5, xmax = 1, ymin = 0, ymax = 1),
    fill = "steelblue", alpha = 0.3
  ) +
  ggplot2::geom_line(
    data = step_df,
    ggplot2::aes(x = x, y = y, group = segment),
    linewidth = 1
  ) +
  ggplot2::geom_point(ggplot2::aes(x = 0.5, y = 0), shape = 1, size = 3) +
  ggplot2::geom_point(ggplot2::aes(x = 0.5, y = 1), shape = 16, size = 3) +
  ggplot2::scale_x_continuous(
    breaks = c(0, 0.5, 1),
    labels = c("0", "1/2", "1")
  ) +
  ggplot2::scale_y_continuous(limits = c(-0.1, 1.2)) +
  ggplot2::labs(x = "x", y = "f(x)") +
  ggplot2::theme_minimal()
```

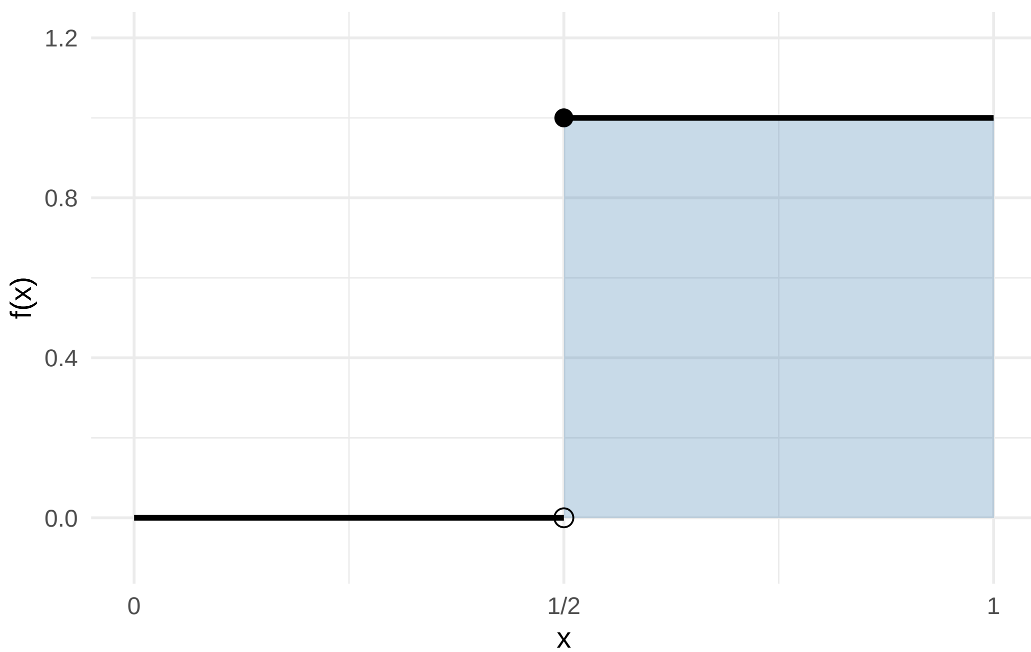


Figure 3: Step function: $f(x) = 0$ on $[0, \frac{1}{2})$ (open circle at the jump) and $f(x) = 1$ on $[\frac{1}{2}, 1]$ (filled circle). The shaded rectangle has area $\frac{1}{2}$, matching the integral computed in Example 0.10.

Together, Theorem 0.10 and Theorem 0.11 establish the chain:

$$\text{differentiable on } [a,b] \;\Rightarrow\; \text{continuous on } [a,b] \;\Rightarrow\; \text{integrable on } [a,b]$$

Example 0.8 and Example 0.10 show that neither implication reverses in general.

Theorem 0.12 (Equal-width Riemann sums converge to the integral)**.** *If f is Riemann integrable on $[a,b]$ (Definition 0.10), then for every choice of the sample points x_i^*, the equal-width Riemann sums (Definition 0.11) converge to the integral:*

$$\lim_{n\to\infty} S_n = \int_a^b f(x)\,dx.$$

Proof

Proof. The n equal-width subintervals form a partition of $[a,b]$ whose mesh (Definition 0.8) is $(b-a)/n$, which goes to 0 as $n \to \infty$. So S_n is one of the sums in the limit that defines the integral (Definition 0.9), along a sequence of partitions whose mesh goes to 0, and a limit that has the same value for every choice of partitions has that value along this sequence too. □

Example 0.11 (Equal-width sums for $\int_0^1 x\,dx$)**.** Let $f(x) = x$ on $[0,1]$, which is continuous and so Riemann integrable (Theorem 0.11), and take each sample point at the right end of its subinterval, $x_i^* = i/n$. With $\Delta x = 1/n$:

$$\begin{aligned}
S_n &= \sum_{i=1}^{n} \frac{i}{n}\cdot\frac{1}{n} && \text{(equal-width Riemann sum with } x_i^* = i/n\text{)}\\
&= \frac{1}{n^2}\sum_{i=1}^{n} i && \text{(factor out } 1/n^2\text{)}\\
&= \frac{1}{n^2}\cdot\frac{n(n+1)}{2} && \text{(sum of the first } n \text{ integers)}\\
&= \frac{n+1}{2n} && \text{(cancel one factor of } n\text{)}
\end{aligned}$$

So $S_{10} = 0.55$, $S_{100} = 0.505$, $S_{1000} = 0.5005$, and $S_n \to \frac{1}{2}$ as $n \to \infty$. By Theorem 0.12, $\int_0^1 x\,dx = \frac{1}{2}$.

Fundamental Theorem of Calculus

Theorem 0.13 (Fundamental Theorem of Calculus)**.** *Let f be a continuous function on a closed interval $[a, b]$.*
Part 1 (Derivative of an integral). *Define $F(x) = \int_a^x f(t)\,dt$ for $x \in [a, b]$. Then F is differentiable and:*

$$\frac{\partial}{\partial x}\int_a^x f(t)\,dt = f(x) \tag{1}$$

> ***Note***
>
> *Continuity on all of $[a, b]$ is a sufficient condition. More generally, Part 1 holds at any individual point x where f is integrable on $[a, b]$ (see Definition 0.10) and continuous at x (see Definition 0.5), even if f has jump discontinuities elsewhere (Rudin 1976, Theorem 6.20, p. 133).*

(Larson and Edwards 2018, Theorem 4.11, p. 288)
Part 2 (Evaluation theorem). *The F here may be* any *antiderivative of f — not just the accumulation function from Part 1. If F is an antiderivative of f on $[a, b]$ (i.e., $\frac{\partial}{\partial x}F(x) = f(x)$ for all $x \in [a, b]$), then:*

$$\int_a^b f(x)\,dx = F(b) - F(a) \tag{2}$$

Equivalently, with b replaced by a variable upper limit x, integrating the derivative of F recovers the net change in F:

$$\int_a^x F'(t)\,dt = F(x) - F(a) \tag{3}$$

or equivalently in Leibniz notation:

$$\int_a^x \frac{dF}{dt}\,dt = F(x) - F(a)$$

(Banner 2007, chap. 18; Larson and Edwards 2018, Theorem 4.9, p. 282)

The two parts of the FTC together express that **differentiation and integration are inverse operations**:

- Part 1: differentiating the integral of f recovers f (Equation 1).
- Part 2: the integral of f over $[a, b]$ equals the difference of any antiderivative's values at the endpoints (Equation 2), which rearranges to "integrating the derivative of F recovers the net change in F" (Equation 3).

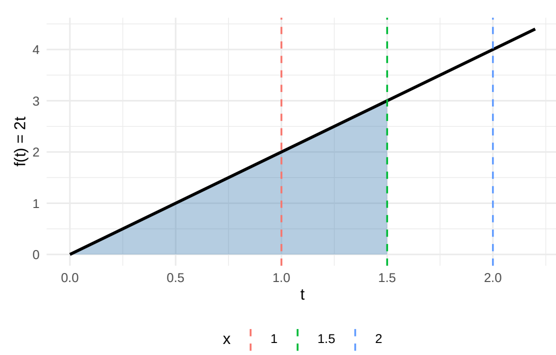
The standard form of the FTC assumes f is continuous on $[a, b]$; continuity is *sufficient* but not strictly necessary (see the callout note inside Theorem 0.13 for the more general statement). Since differentiability implies continuity (Theorem 0.10), the FTC applies in particular whenever f is differentiable — a common situation in applied statistics.

Example 0.12 (FTC Part 1 visualized: accumulation function for $f(t) = 2t$)**.** Take $f(t) = 2t$ on $[0, 2]$. The accumulation function from 0 is

$$F(x) \stackrel{\text{def}}{=} \int_0^x 2t\,dt \;=\; [t^2]_{t=0}^{t=x} \;=\; x^2 - 0^2 \;=\; x^2,$$

so $F(x) = x^2$, and indeed $F'(x) = 2x = f(x)$, as Theorem 0.13 Part 1 predicts. Figure 4 shows the integrand on the left (shaded area equals $F(x)$ at each x) and the accumulation function $F(x) = x^2$ on the right (its slope at x equals $f(x) = 2x$).

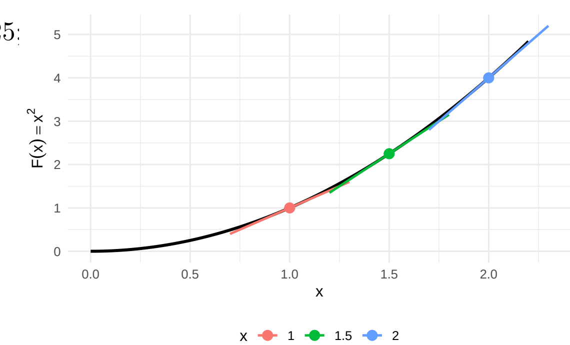
```r
ggplot2::ggplot() +
  ggplot2::geom_area(
    data = data.frame(t = seq(0, x_f
    ggplot2::aes(x = t, y = 2 * t),
    fill = "steelblue", alpha = 0.4
  ) +
  ggplot2::geom_function(fun = \(t) ) +
  ggplot2::geom_vline(
    data = data.frame(x = x_marks),
    ggplot2::aes(xintercept = x, col
    linetype = "dashed", linewidth =
  ) +
  ggplot2::labs(x = "t", y = "f(t) =
  ggplot2::theme_minimal() +
  ggplot2::theme(legend.position = "
```



(a) $f(t) = 2t$; shaded area equals $F(1.5) = 2.25$; vertical lines mark $x \in \{1, 1.5, 2\}$.

```r
slope_df <- data.frame(
  x = x_marks,
  Fx = x_marks^2,
  slope = 2 * x_marks
)

ggplot2::ggplot() +
  ggplot2::geom_function(fun = \(x) x^2, xlim = c(0, 2
  ggplot2::geom_point(
    data = slope_df,
    ggplot2::aes(x = x, y = Fx, color = factor(x)),
    size = 3
  ) +
  ggplot2::geom_segment(
    data = slope_df,
    ggplot2::aes(
      x = x - 0.3, xend = x + 0.3,
      y = Fx - 0.3 * slope, yend = Fx + 0.3 * slope,
      color = factor(x)
    ),
    linewidth = 0.8
  ) +
  ggplot2::labs(x = "x", y = expression(F(x) == x^2),
  ggplot2::theme_minimal() +
  ggplot2::theme(legend.position = "bottom")
```



(b) $F(x) = x^2$; tangent slope at each marked x equals $f(x) = 2x$.

Figure 4: Left: $f(t) = 2t$; the shaded area $\int_0^{1.5} 2t\,dt = F(1.5) = 2.25$; vertical lines mark $x \in \{1, 1.5, 2\}$. Right: $F(x) = x^2$; for each marked x, the tangent slope equals $f(x) = 2x$.

Example 0.13 (CDF and PDF of the exponential distribution)**.** In what follows, f denotes the PDF and F the CDF — the same letters as the antiderivative pair in Definition 0.1, because the FTC will show F is exactly an antiderivative of f.
Let T be a random variable with the exponential distribution with rate parameter $\lambda > 0$. Its probability density function (PDF) is (Kleinbaum and Klein 2012, sec. II, p. 295, "Survival and Hazard Functions for Selected Distributions"):

$$f(t) = \lambda \mathrm{e}^{-\lambda t}, \quad t \geq 0$$

FTC Part 2 gives the cumulative distribution function (CDF), $F(t) = P(T \leq t)$, from the PDF. Apply the e^{cx} rule from Theorem 0.9 with $c = -\lambda$ to antidifferentiate the integrand:

$$\begin{aligned}
F(t) &= \int_0^t \lambda \mathrm{e}^{-\lambda u}\, du && \text{(the CDF integrates the PDF)} \\
&= \left[\lambda \cdot \frac{1}{-\lambda} \mathrm{e}^{-\lambda u}\right]_{u=0}^{u=t} && \text{(FTC Part 2, with the } \mathrm{e}^{cx} \text{ rule)} \\
&= \left[(-1)\mathrm{e}^{-\lambda u}\right]_{u=0}^{u=t} && (\lambda/(-\lambda) = -1) \\
&= \left[-\mathrm{e}^{-\lambda u}\right]_{u=0}^{u=t} && \text{(multiply by } -1) \\
&= -\mathrm{e}^{-\lambda t} - (-\mathrm{e}^{0}) && \text{(evaluate at the limits)} \\
&= -\mathrm{e}^{-\lambda t} - (-1) && (\mathrm{e}^0 = 1) \\
&= 1 - \mathrm{e}^{-\lambda t} && \text{(rearrange)}
\end{aligned}$$

FTC Part 1 recovers the PDF from the CDF:

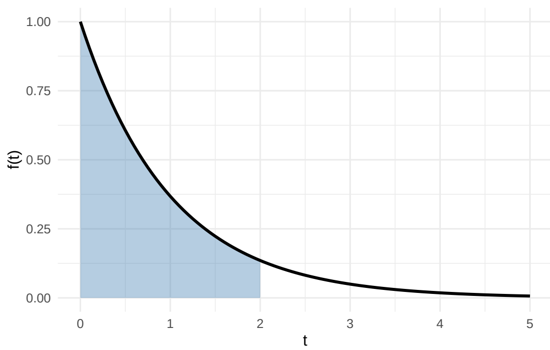
$$\begin{aligned}
\frac{\partial}{\partial t} F(t) &= \frac{\partial}{\partial t}(1 - \mathrm{e}^{-\lambda t}) && \text{(substitute } F) \\
&= \frac{\partial}{\partial t} 1 - \frac{\partial}{\partial t} \mathrm{e}^{-\lambda t} && \text{(derivative of a difference)} \\
&= 0 - \frac{\partial}{\partial t} \mathrm{e}^{-\lambda t} && \text{(constant rule)} \\
&= 0 - \mathrm{e}^{-\lambda t} \cdot \frac{\partial}{\partial t}(-\lambda t) && \text{(chain rule, with inner function } -\lambda t) \\
&= 0 - \mathrm{e}^{-\lambda t} \cdot (-\lambda) && \text{(constant multiple rule)} \\
&= \lambda \mathrm{e}^{-\lambda t} && \text{(simplify)} \\
&= f(t) && \text{(definition of } f)
\end{aligned}$$

For a concrete instance: with $\lambda = 1$ (standard exponential), the probability that $T \leq 2$ is:

$$F(2) = 1 - \mathrm{e}^{-1 \cdot 2} = 1 - \mathrm{e}^{-2} \approx 1 - 0.135 = 0.865$$

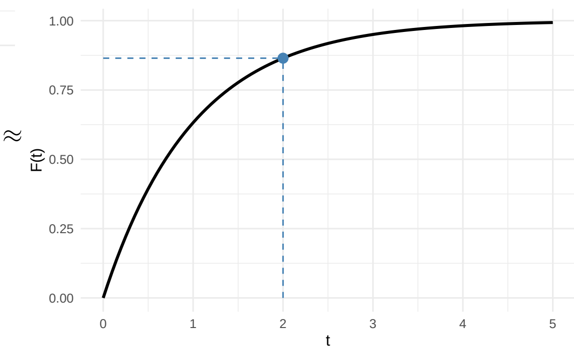
See Figure 5.

```r
ggplot2::ggplot() +
  ggplot2::geom_area(
    data = data.frame(t = seq(0, t_f
    ggplot2::aes(x = t, y = lambda *
    fill = "steelblue", alpha = 0.4
  ) +
  ggplot2::geom_function(
    fun = \(t) lambda * exp(-lambda
    xlim = c(0, t_max), linewidth =
  ) +
  ggplot2::labs(x = "t", y = "f(t)")
  ggplot2::theme_minimal()
```



(a) PDF with $\lambda = 1$; shaded area equals $F(2) \approx 0.865$.

```r
ggplot2::ggplot() +
  ggplot2::geom_function(
    fun = \(t) 1 - exp(-lambda * t),
    xlim = c(0, t_max), linewidth = 1
  ) +
  ggplot2::geom_point(
    ggplot2::aes(x = t_focus, y = F_at_focus),
    size = 3, color = "steelblue"
  ) +
  ggplot2::geom_segment(
    ggplot2::aes(x = t_focus, xend = t_focus, y = 0, y
    linetype = "dashed", color = "steelblue"
  ) +
  ggplot2::geom_segment(
    ggplot2::aes(x = 0, xend = t_focus, y = F_at_focus
    linetype = "dashed", color = "steelblue"
  ) +
  ggplot2::labs(x = "t", y = "F(t)") +
  ggplot2::theme_minimal()
```



(b) CDF with $\lambda = 1$; point marks $F(2) \approx 0.865$.

Figure 5: Exponential distribution with $\lambda = 1$. Left: the PDF $f(t) = \lambda e^{-\lambda t}$; the shaded area under the curve from 0 to 2 equals $F(2) \approx 0.865$. Right: the CDF $F(t) = 1 - e^{-\lambda t}$; the dashed lines mark the value $F(2)$ computed via FTC Part 2.

Double Integrals

The **Fubini–Tonelli theorem** states conditions under which the order of integration in a double integral can be exchanged. We state two versions: the Riemann version (Theorem 0.14) is what applied courses usually use for double integrals of continuous functions on simple regions; the σ-finite measure-theoretic version (Theorem 0.15) is included to make the joint-distribution form corollary in the probability chapter of *Regression Models for Epidemiology* follow from a stated theorem rather than from an aside.

Definition 0.12 (Double integral)**.** Let f be a bounded function on a closed, bounded plane region $R \subseteq \mathbb{R}^2$. Cover R with a grid of rectangles, keep the n rectangles that lie entirely inside R, with areas $\Delta A_1, \dots, \Delta A_n$, and choose a point (x_i, y_i) in the i-th rectangle. The **double integral** of f over R is

$$\iint_R f(x,y)\,dA \overset{\text{def}}{=} \lim_{\|\Delta\| \to 0} \sum_{i=1}^{n} f(x_i, y_i)\,\Delta A_i,$$

where $\|\Delta\|$ is the length of the longest diagonal among the n rectangles, when that limit exists and has the same value for every choice of grids and of the points (x_i, y_i).
(Larson and Edwards 2018, sec. 14.2)

The symbol dA stands for an element of area. An *iterated integral* such as $\int_a^b \int_c^d f(x,y)\,dy\,dx$ means $\int_a^b \left(\int_c^d f(x,y)\,dy\right) dx$: integrate over the inner variable (y) first, holding the outer variable (x) fixed.

Example 0.14 (The double integral of 1 is an area)**.** Let $f(x,y) = 1$ on the rectangle $R = [0,2] \times [0,3]$. Every sum in Definition 0.12 adds up the areas of rectangles inside R, and those sums approach the area of R as the grid gets finer, so $\iint_R 1\,dA = 2 \cdot 3 = 6$.

Theorem 0.14 (Fubini's theorem (Riemann version))**.** *Let f be **continuous** on a plane region $R \subseteq \mathbb{R}^2$.*

1. ***Vertically simple region.*** *If R is defined by $a \le x \le b$ and $g_1(x) \le y \le g_2(x)$, where g_1 and g_2 are continuous on $[a,b]$, then*

$$\iint_R f(x,y)\,dA = \int_a^b \int_{g_1(x)}^{g_2(x)} f(x,y)\,dy\,dx.$$

2. ***Horizontally simple region.*** *If R is defined by $c \le y \le d$ and $h_1(y) \le x \le h_2(y)$, where h_1 and h_2 are continuous on $[c,d]$, then*

$$\iint_R f(x,y)\,dA = \int_c^d \int_{h_1(y)}^{h_2(y)} f(x,y)\,dx\,dy.$$

When R can be described both ways, the two iterated integrals are equal — so the order of integration can be exchanged.

(Larson and Edwards 2018, Theorem 14.2, p. 982)

Example 0.15 (Changing the order of integration for a non-rectangular region)**.** Adapted from (Larson and Edwards 2018, sec. 14.2, Example 4, pp. 984–985).
Let X and Y be independent Uniform$(0, 1)$ random variables, with joint density $f(x, y) = 1$ on the unit square $[0, 1]^2$. Define the function $g(x, y) = \mathrm{e}^{-x^2}\, \mathbb{1}(y \le x)$, and compute its expectation $\mathrm{E}[g(X, Y)]$.
Because the joint density equals 1 on $[0, 1]^2$, this expectation is the double integral of g over the unit square:

$$\mathrm{E}[g(X, Y)] = \iint_{[0,1]^2} g(x, y)\, dA.$$

The indicator factor $\mathbb{1}(y \le x)$ equals 1 on the triangular region where $y \le x$ and 0 elsewhere, so only that region, namely $D = \{(x, y) : x \in [0, 1],\ y \in [0, x]\}$ (Figure 6), contributes, and there $g(x, y) = \mathrm{e}^{-x^2}$:

$$\mathrm{E}[g(X, Y)] = \iint_D \mathrm{e}^{-x^2}\, dA.$$

```r
region <- data.frame(x = c(0, 1, 1), y = c(0, 0, 1))
ggplot2::ggplot(region, ggplot2::aes(x = x, y = y)) +
  ggplot2::geom_polygon(
    fill = "steelblue", alpha = 0.4, color = "black", linewidth = 0.7
  ) +
  ggplot2::annotate(
    "text", x = 0.65, y = 0.28, label = "D", size = 8, fontface = "italic"
  ) +
  ggplot2::annotate(
    "text", x = 0.36, y = 0.52, label = "y == x",
    parse = TRUE, angle = 45
  ) +
  ggplot2::annotate(
    "text", x = 1.08, y = 0.5, label = "x == 1",
    parse = TRUE, angle = 90, hjust = 0.5
  ) +
  ggplot2::annotate(
    "text", x = 0.5, y = -0.1, label = "y == 0", parse = TRUE
  ) +
  ggplot2::scale_x_continuous(breaks = c(0, 0.5, 1), limits = c(-0.1, 1.3)) +
  ggplot2::scale_y_continuous(breaks = c(0, 0.5, 1), limits = c(-0.2, 1.2)) +
  ggplot2::coord_equal() +
  ggplot2::labs(x = "x", y = "y") +
  ggplot2::theme_minimal()
```

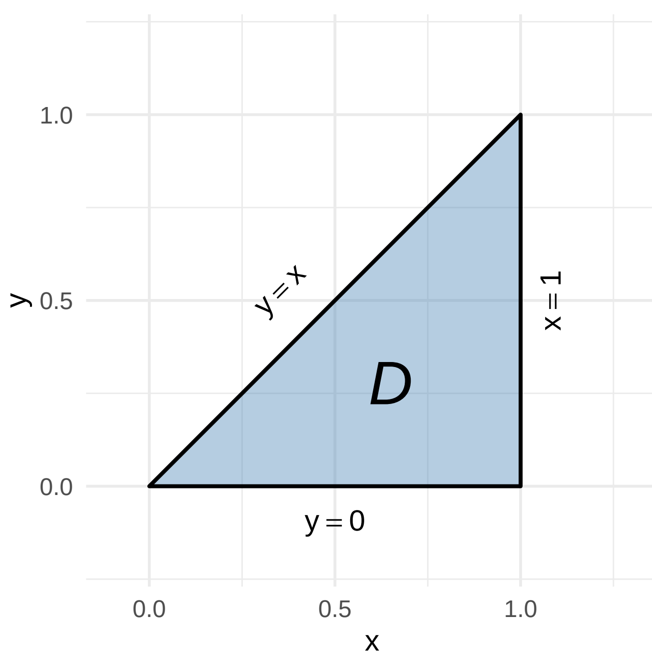


Figure 6: Triangular integration region $D = \{(x, y) : x \in [0, 1],\ y \in [0, x]\}$, bounded below by $y = 0$, above-left by $y = x$, and on the right by $x = 1$.

Order $dx\,dy$ is intractable. Re-describing D as $D = \{(x,y) : y \in [0,1],\ x \in [y,1]\}$, the inner integral is

$$\int_y^1 \mathrm{e}^{-x^2}\,dx,$$

which has no elementary antiderivative.

Order $dy\,dx$ works. Applying Theorem 0.14 Part 1 (e^{-x^2} is continuous and D is the vertically simple region $x \in [0,1]$, $y \in [0,x]$):

$$\begin{aligned}
\iint_D \mathrm{e}^{-x^2}\,dA &= \int_0^1 \int_0^x \mathrm{e}^{-x^2}\,dy\,dx && \text{(Fubini, vertically simple region)}\\
&= \int_0^1 \mathrm{e}^{-x^2}\left(\int_0^x dy\right) dx && (\mathrm{e}^{-x^2} \text{ is constant in } y)\\
&= \int_0^1 x\,\mathrm{e}^{-x^2}\,dx && \left(\textstyle\int_0^x dy = x\right)
\end{aligned}$$

To antidifferentiate $x\,\mathrm{e}^{-x^2}$, let $u = x^2$ be the inner function. By the chain rule (Theorem 0.8), with $\frac{du}{dx} = 2x$,

$$\begin{aligned}
\frac{d}{dx}\left(-\frac{1}{2}\,\mathrm{e}^{-u}\right) &= -\frac{1}{2}\,\mathrm{e}^{-u}\cdot(-1)\cdot\frac{du}{dx} && \text{(chain rule)}\\
&= -\frac{1}{2}\,\mathrm{e}^{-x^2}\cdot(-1)\cdot 2x && (\text{substitute } u = x^2 \text{ and } du/dx = 2x)\\
&= x\,\mathrm{e}^{-x^2} && \text{(multiply)}
\end{aligned}$$

so $-\frac{1}{2}\,\mathrm{e}^{-x^2}$ is an antiderivative of $x\,\mathrm{e}^{-x^2}$, and

$$\begin{aligned}
\int_0^1 x\,\mathrm{e}^{-x^2}\,dx &= \left[-\frac{1}{2}\,\mathrm{e}^{-x^2}\right]_0^1 && \text{(FTC Part 2)}\\
&= -\frac{1}{2}(\mathrm{e}^{-1} - \mathrm{e}^{0}) && \text{(evaluate at the limits)}\\
&= -\frac{1}{2}(\mathrm{e}^{-1} - 1) && (\mathrm{e}^0 = 1)\\
&= \frac{1 - \mathrm{e}^{-1}}{2} && \left(\text{distribute } -\tfrac{1}{2}\right)\\
&= \frac{e-1}{2e} && (\text{multiply numerator and denominator by } e)\\
&\approx 0.316
\end{aligned}$$

The solid whose volume equals this integral is shown in Figure 7.

(This 3D surface is interactive; see the HTML or slide version of this page to view it.)

Figure 7: Surface $z = e^{-x^2}$ over the region $D = \{(x, y) : x \in [0, 1],\ y \in [0, x]\}$. The surface depends only on x (constant in y), so for each x the inner integral over $y \in [0, x]$ contributes $x \cdot e^{-x^2}$.

Example 0.16 (When conditions fail: a counterexample)**.** The conditions in Theorem 0.14 are not merely technical — when they fail, iterated integrals can exist yet disagree.
Let

$$f(x, y) = \frac{x^2 - y^2}{(x^2 + y^2)^2}$$

on the unit square $R = [0, 1] \times [0, 1]$. Strictly, f is defined on $R \setminus \{(0, 0)\}$: the denominator vanishes at the origin, so f is undefined there (we return to this point in the condition check).

Integrating y first, then x:
Using $\dfrac{\partial}{\partial y} \dfrac{y}{x^2 + y^2} = \dfrac{x^2 - y^2}{(x^2 + y^2)^2}$:

$$\begin{aligned} \int_0^1 \int_0^1 f(x, y)\, dy\, dx &= \int_0^1 \left[\frac{y}{x^2 + y^2}\right]_{y=0}^{y=1} dx \\ &= \int_0^1 \frac{1}{x^2 + 1}\, dx \\ &= [\arctan(x)]_0^1 \\ &= \frac{\pi}{4} \end{aligned}$$

Integrating x first, then y:
Using $\dfrac{\partial}{\partial x} \left(-\dfrac{x}{x^2 + y^2}\right) = \dfrac{x^2 - y^2}{(x^2 + y^2)^2}$:

$$\begin{aligned} \int_0^1 \int_0^1 f(x, y)\, dx\, dy &= \int_0^1 \left[-\frac{x}{x^2 + y^2}\right]_{x=0}^{x=1} dy \\ &= \int_0^1 \left(-\frac{1}{1 + y^2}\right) dy \\ &= -[\arctan(y)]_0^1 \\ &= -\frac{\pi}{4} \end{aligned}$$

Conclusion: $\dfrac{\pi}{4} \neq -\dfrac{\pi}{4}$, so the two iterated integrals are unequal. Theorem 0.14 does not apply here.
Why Theorem 0.14's condition fails: Theorem 0.14 requires f to be **continuous** on R. The denominator $(x^2 + y^2)^2$ vanishes at the origin $(0, 0) \in R$, so f is *not even defined*

there — let alone continuous — and the theorem does not apply.
(Wikipedia contributors 2024)
The surface, and the singularity at the origin responsible for the failure, are shown in Figure 8.

(This 3D surface is interactive; see the HTML or slide version of this page to view it.)

Figure 8: Surface $f(x, y) = (x^2 - y^2)/(x^2 + y^2)^2$ on $[0, 1]^2$, sampled away from the origin and clipped to $[-50, 50]$ for display. The function diverges to $+\infty$ along the x-axis (red ridge, $f > 0$ when $|x| > |y|$) and to $-\infty$ along the y-axis (blue ridge, $f < 0$ when $|y| > |x|$). The singularity at $(0, 0)$ is why f is not continuous on R and Theorem 0.14 does not apply.

Corollary 0.2 (Continuous functions on a rectangle (corollary of Theorem 0.14))**.** *If* $f : [a, b] \times [c, d] \to \mathbb{R}$ *is **continuous** on the closed bounded rectangle* $[a, b] \times [c, d]$*, then:*

$$\int_a^b \left(\int_c^d f(x, y)\, dy \right) dx = \int_c^d \left(\int_a^b f(x, y)\, dx \right) dy$$
$$= \iint_{[a,b]\times[c,d]} f(x, y)\, dA.$$

(Larson and Edwards 2018, Theorem 14.2, p. 982)

Proof

Proof. A closed bounded rectangle $[a, b] \times [c, d]$ is both vertically simple (with $g_1 \equiv c$, $g_2 \equiv d$) and horizontally simple (with $h_1 \equiv a$, $h_2 \equiv b$). Applying both parts of Theorem 0.14 to f on this rectangle gives the two iterated forms shown. □

Example 0.17 (Evaluating a double integral on a rectangle)**.** Structure adapted from (Larson and Edwards 2018, sec. 14.2, Example 2, pp. 982–983); the integrand $x^2 + y^2$ is original, chosen so the integral equals $\mathrm{E}[g(X, Y)]$ for $g(x, y) = x^2 + y^2$.
Let X and Y be independent Uniform$(0, 1)$ random variables, with joint density $f(x, y) = 1$ on the unit square $R = \{(x, y) : x \in [0, 1],\ y \in [0, 1]\}$ (Figure 9). Define the function $g(x, y) = x^2 + y^2$, and compute its expectation $\mathrm{E}[g(X, Y)]$.
Because the joint density equals 1 on R, this expectation is the double integral of g over R:

$$\mathrm{E}[g(X, Y)] = \iint_R (x^2 + y^2)\, dA.$$

```r
ggplot2::ggplot() +
  ggplot2::annotate(
    "rect", xmin = 0, xmax = 1, ymin = 0, ymax = 1,
    fill = "steelblue", alpha = 0.4, color = "black", linewidth = 0.7
  ) +
  ggplot2::annotate(
    "text", x = 0.5, y = 0.5, label = "R", size = 8, fontface = "italic"
  ) +
  ggplot2::scale_x_continuous(breaks = c(0, 1), limits = c(-0.15, 1.25)) +
  ggplot2::scale_y_continuous(breaks = c(0, 1), limits = c(-0.15, 1.25)) +
  ggplot2::coord_equal() +
  ggplot2::labs(x = "x", y = "y") +
  ggplot2::theme_minimal()
```

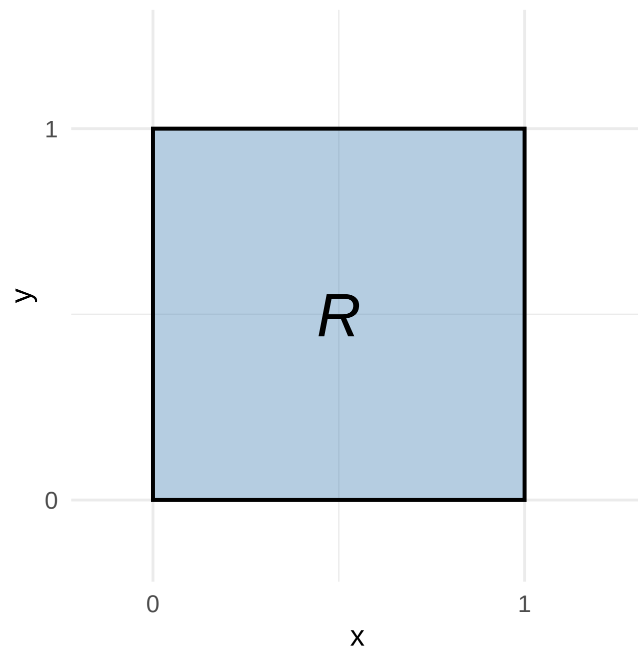


Figure 9: Integration region $R = [0, 1]^2$, the unit square.

The integrand is continuous on R, so Corollary 0.2 applies and either order of integration yields the same value.

Integrating y first, then x:

$$
\begin{aligned}
\mathrm{E}[g(X,Y)] &= \int_0^1 \int_0^1 (x^2+y^2)\,dy\,dx \\
&= \int_0^1 \left[x^2 y + \frac{y^3}{3}\right]_0^1 dx \\
&= \int_0^1 \left(x^2 + \frac{1}{3}\right) dx \\
&= \left[\frac{x^3}{3} + \frac{x}{3}\right]_0^1 \\
&= \frac{2}{3}
\end{aligned}
$$

Integrating x first, then y (verifying the order can be swapped):

$$
\begin{aligned}
\int_0^1 \int_0^1 (x^2+y^2)\,dx\,dy &= \int_0^1 \left[\frac{x^3}{3} + y^2 x\right]_0^1 dy \\
&= \int_0^1 \left(\frac{1}{3} + y^2\right) dy \\
&= \left[\frac{y}{3} + \frac{y^3}{3}\right]_0^1 \\
&= \frac{2}{3}
\end{aligned}
$$

Both orders give $\frac{2}{3}$, as Corollary 0.2 guarantees.

As a cross-check, linearity of expectation gives the same value: since $\mathrm{E}[X^2] = \int_0^1 x^2\,dx = \frac{1}{3}$ for $X \sim \mathrm{Uniform}(0,1)$ (and likewise for Y),

$$
\mathrm{E}[g(X,Y)] = \mathrm{E}[X^2+Y^2] = \mathrm{E}[X^2] + \mathrm{E}[Y^2] = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}.
$$

The solid whose volume equals this integral is shown in Figure 10.

(This 3D surface is interactive; see the HTML or slide version of this page to view it.)

Figure 10: Surface $z = x^2 + y^2$ over the unit square $[0,1]^2$. The double integral $\frac{2}{3}$ is the volume between this surface and the xy-plane, and equals $\mathrm{E}[X^2+Y^2]$.

Theorem 0.15 (Fubini–Tonelli theorem (measure-theoretic form))**.** *Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ be **σ-finite** measure spaces, and let $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be measurable with respect to the product σ-algebra $\mathcal{F}_1 \otimes \mathcal{F}_2$. If either*

*(a) $f \geq 0$ almost everywhere with respect to $\mu_1 \otimes \mu_2$ (**Tonelli's theorem**), or*

(b) $\int_{\Omega_1 \times \Omega_2} |f| \, d(\mu_1 \otimes \mu_2) < \infty$ ***(Fubini's theorem)****,*

then both iterated integrals exist, agree with the double integral, and equal each other:

$$\begin{aligned}\int_{\Omega_1 \times \Omega_2} f \, d(\mu_1 \otimes \mu_2) &= \int_{\Omega_1} \left(\int_{\Omega_2} f(\omega_1, \omega_2) \, d\mu_2(\omega_2) \right) d\mu_1(\omega_1) \\ &= \int_{\Omega_2} \left(\int_{\Omega_1} f(\omega_1, \omega_2) \, d\mu_1(\omega_1) \right) d\mu_2(\omega_2).\end{aligned}$$

(Billingsley 1995, Theorem 18.3; Gut 2013, Theorem 9.1, p. 65; Fubini 1907; Wikipedia contributors 2024)

Applied courses rarely need the measure-theoretic generalization itself, but it is what justifies the joint-distribution form corollary in the probability chapter of *Regression Models for Epidemiology*: probability measures are finite (hence σ-finite), so the σ-finiteness condition is automatic. The integrability conditions (nonnegativity or absolute integrability) still need to be verified in each application.

Example 0.18 (Positive application of Theorem 0.15)**.** Let X and Y be independent Exponential(1) random variables, with joint density $f(x, y) = e^{-(x+y)}$ for $x, y \geq 0$.
The probability $P(X \leq 1, Y \leq 1)$ is the integral of f over $[0, 1]^2$ with respect to Lebesgue measure (ordinary length) in each coordinate. Lebesgue measure on $[0, \infty)$ is σ-finite, because $[0, \infty)$ is the union of the intervals $[0, n]$, $n \in \mathbb{N}$, each of finite length n; so the σ-finiteness condition of Theorem 0.15 holds. Since $f(x, y) = e^{-(x+y)} \geq 0$, condition (a) (Tonelli's theorem, nonnegativity) is also satisfied.
By Theorem 0.15, both iterated integrals exist and agree. Integrating y first, then x:

$$
\begin{aligned}
P(X \le 1, Y \le 1) &= \int_0^1 \int_0^1 e^{-(x+y)}\,dy\,dx \\
&= \int_0^1 \int_0^1 e^{-x}e^{-y}\,dy\,dx && \text{(exponential of a sum)} \\
&= \int_0^1 e^{-x} \left(\int_0^1 e^{-y}\,dy \right) dx && (e^{-x} \text{ is constant in } y) \\
&= \int_0^1 e^{-x} [-e^{-y}]_{y=0}^{y=1}\,dx && (\text{antiderivative of } e^{-y}) \\
&= \int_0^1 e^{-x}(1-e^{-1})\,dx && \text{(evaluate at the limits)} \\
&= (1-e^{-1}) \int_0^1 e^{-x}\,dx && (1-e^{-1} \text{ is constant in } x) \\
&= (1-e^{-1})[-e^{-x}]_{x=0}^{x=1} && (\text{antiderivative of } e^{-x}) \\
&= (1-e^{-1})^2 && \text{(evaluate at the limits)}
\end{aligned}
$$

Integrating x first, then y:

$$
\begin{aligned}
P(X \le 1, Y \le 1) &= \int_0^1 \int_0^1 e^{-(x+y)}\,dx\,dy \\
&= \int_0^1 \int_0^1 e^{-y}e^{-x}\,dx\,dy && \text{(exponential of a sum)} \\
&= \int_0^1 e^{-y} \left(\int_0^1 e^{-x}\,dx \right) dy && (e^{-y} \text{ is constant in } x) \\
&= \int_0^1 e^{-y} [-e^{-x}]_{x=0}^{x=1}\,dy && (\text{antiderivative of } e^{-x}) \\
&= \int_0^1 e^{-y}(1-e^{-1})\,dy && \text{(evaluate at the limits)} \\
&= (1-e^{-1}) \int_0^1 e^{-y}\,dy && (1-e^{-1} \text{ is constant in } y) \\
&= (1-e^{-1})[-e^{-y}]_{y=0}^{y=1} && (\text{antiderivative of } e^{-y}) \\
&= (1-e^{-1})^2 && \text{(evaluate at the limits)}
\end{aligned}
$$

Both iterated integrals equal $(1-e^{-1})^2 \approx 0.400$, as Theorem 0.15 guarantees when condition (a) holds.

Example 0.19 (When neither Fubini–Tonelli condition is satisfied)**.** The same function $f(x,y) = (x^2-y^2)/(x^2+y^2)^2$ from Example 0.16 illustrates a case where neither condition of Theorem 0.15 is satisfied.

Why Theorem 0.15's conditions fail: $\iint_R |f|\,dA = \infty$, which violates condition (b). Switching to polar coordinates (r,θ) near the origin, the integrand satisfies $|f(x,y)| = |x^2-y^2|/(x^2+y^2)^2 = |\cos 2\theta|/r^2$, so

$$\begin{aligned}\iint_R |f|\,dA &\geq \int_0^{\pi/2}\int_0^{\epsilon} \frac{|\cos 2\theta|}{r^2}\, r\,dr\,d\theta \\ &= \left(\int_0^{\pi/2} |\cos 2\theta|\,d\theta\right)\int_0^{\epsilon}\frac{dr}{r} \\ &= +\infty,\end{aligned}$$

since $\int_0^{\epsilon} dr/r$ diverges. Therefore $\iint_R |f|\,dA = \infty$, and condition (b) of Theorem 0.15 is not satisfied. (Condition (a) also fails: f takes both positive and negative values, so it is not nonnegative a.e.) The unequal iterated integrals from Example 0.16 are thus consistent with Theorem 0.15: the theorem simply does not apply.
(Wikipedia contributors 2024)

Additional resources

- Kaplan (2022)
- Khuri (2003)
- Banner (2007)
- Larson and Edwards (2018)
- Miller (2016)
  - http://www.youtube.com/watch?v=xYzQL0TUtBA
  - http://www.youtube.com/watch?v=Ps2SBo_WjoE
- Grinberg (2017) (the rigorous foundations behind these results)

References

Banner, Adrian D. 2007. *The Calculus Lifesaver : All the Tools You Need to Excel at Calculus*. A Princeton Lifesaver Study Guide. Princeton University Press. https://press.princeton.edu/books/paperback/9780691130880/the-calculus-lifesaver.

Billingsley, Patrick. 1995. *Probability and Measure*. 3rd ed. Wiley Series in Probability and Mathematical Statistics. Wiley.

Fubini, Guido. 1907. "Sugli Integrali Multipli." *Rendiconti Della Reale Accademia Dei Lincei. Classe Di Scienze Fisiche, Matematiche e Naturali* 16: 608–14.

Grinberg, Raffi. 2017. *The Real Analysis Lifesaver: All the Tools You Need to Understand Proofs*. 1st ed. Princeton Lifesaver Study Guides. Princeton University Press. https://press.princeton.edu/books/paperback/9780691172934/the-real-analysis-lifesaver.

Gut, Allan. 2013. *Probability: A Graduate Course*. 2nd ed. Springer Texts in Statistics. Springer. https://doi.org/10.1007/978-1-4614-4708-5.

Kaplan, Daniel. 2022. *MOSAIC Calculus*. Www.mosaic-web.org. www.mosaic-web.org.

Khuri, André I. 2003. *Advanced Calculus with Applications in Statistics*. John Wiley & Sons. https://doi.org/10.1002/0471394882.

Kleinbaum, David G, and Mitchel Klein. 2012. *Survival Analysis: A Self-Learning Text*. 3rd ed. Springer. https://doi.org/10.1007/978-1-4419-6646-9.

Larson, Ron, and Bruce H. Edwards. 2018. *Calculus*. 11th ed. Cengage Learning. https://www.cengage.com/c/calculus-11e-larson/.

Miller, Steven J. 2016. *The Probability Lifesaver: Calculus Review Problems*. https://web.williams.edu/Mathematics/sjmiller/public_html/probabilitylifesaver/index.htm#:~:text=http%3A//web.williams.edu/Mathematics/sjmiller/public_html/probabilitylifesaver/supplementalchap_calcreview.pdf.

Rudin, Walter. 1976. *Principles of Mathematical Analysis*. 3rd ed. International Series in Pure and Applied Mathematics. McGraw-Hill.

Wikipedia contributors. 2024. *Fubini's Theorem — Wikipedia, the Free Encyclopedia*. https://en.wikipedia.org/wiki/Fubini%27s_theorem.