18.440: Lecture 17

Continuous random variables

Scott Sheffield

MIT

Outline

Outline

Continuous random variables

- Say X is a **continuous random variable** if there exists a **probability density function** $f = f_X$ on $\mathbb{R}$ such that $P\{X \in B\} = \int_B f(x)dx := \int 1_B(x)f(x)dx$.

Continuous random variables

- Say X is a **continuous random variable** if there exists a **probability density function** $f = f_X$ on $\mathbb{R}$ such that $P\{X \in B\} = \int_B f(x)dx := \int 1_B(x)f(x)dx$.
- We may assume $\int_{\mathbb{R}} f(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1$ and f is non-negative.

Continuous random variables

- Say X is a **continuous random variable** if there exists a **probability density function** $f = f_X$ on $\mathbb{R}$ such that $P\{X \in B\} = \int_B f(x)dx := \int 1_B(x)f(x)dx$.
- We may assume $\int_{\mathbb{R}} f(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1$ and f is non-negative.
- Probability of interval $[a, b]$ is given by $\int_a^b f(x)dx$, the area under f between a and b.

Continuous random variables

- Say X is a **continuous random variable** if there exists a **probability density function** $f = f_X$ on $\mathbb{R}$ such that $P\{X \in B\} = \int_B f(x)dx := \int 1_B(x)f(x)dx$.
- We may assume $\int_{\mathbb{R}} f(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1$ and f is non-negative.
- Probability of interval $[a, b]$ is given by $\int_a^b f(x)dx$, the area under f between a and b.
- Probability of any single point is zero.

Continuous random variables

- Say X is a **continuous random variable** if there exists a **probability density function** $f = f_X$ on $\mathbb{R}$ such that $P\{X \in B\} = \int_B f(x)dx := \int 1_B(x)f(x)dx$.
- We may assume $\int_{\mathbb{R}} f(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1$ and f is non-negative.
- Probability of interval $[a, b]$ is given by $\int_a^b f(x)dx$, the area under f between a and b.
- Probability of any single point is zero.
- Define **cumulative distribution function** $F(a) = F_X(a) := P\{X < a\} = P\{X \leq a\} = \int_{-\infty}^{a} f(x)dx$.

Simple example

- Suppose $f(x) = \begin{cases} 1/2 & x \in [0,2] \\ 0 & x \notin [0,2]. \end{cases}$

Simple example

- Suppose $f(x) = \begin{cases} 1/2 & x \in [0,2] \\ 0 & x \notin [0,2]. \end{cases}$
- What is $P\{X < 3/2\}$?

Simple example

- Suppose $f(x) = \begin{cases} 1/2 & x \in [0,2] \\ 0 & x \notin [0,2]. \end{cases}$
- What is $P\{X < 3/2\}$?
- What is $P\{X = 3/2\}$?

Simple example

- Suppose $f(x) = \begin{cases} 1/2 & x \in [0,2] \\ 0 & x \notin [0,2]. \end{cases}$
- What is $P\{X < 3/2\}$?
- What is $P\{X = 3/2\}$?
- What is $P\{1/2 < X < 3/2\}$?

Simple example

- Suppose $f(x) = \begin{cases} 1/2 & x \in [0,2] \\ 0 & x \notin [0,2]. \end{cases}$
- What is $P\{X < 3/2\}$?
- What is $P\{X = 3/2\}$?
- What is $P\{1/2 < X < 3/2\}$?
- What is $P\{X \in (0,1) \cup (3/2, 5)\}$?

Simple example

- Suppose $f(x) = \begin{cases} 1/2 & x \in [0,2] \\ 0 & x \notin [0,2]. \end{cases}$
- What is $P\{X < 3/2\}$?
- What is $P\{X = 3/2\}$?
- What is $P\{1/2 < X < 3/2\}$?
- What is $P\{X \in (0,1) \cup (3/2,5)\}$?
- What is F?

Simple example

- Suppose $f(x) = \begin{cases} 1/2 & x \in [0,2] \\ 0 & x \notin [0,2]. \end{cases}$
- What is $P\{X < 3/2\}$?
- What is $P\{X = 3/2\}$?
- What is $P\{1/2 < X < 3/2\}$?
- What is $P\{X \in (0,1) \cup (3/2,5)\}$?
- What is F?
- We say that X is **uniformly distributed on the interval** $[0,2]$.

Another example

- Suppose $f(x) = \begin{cases} x/2 & x \in [0,2] \\ 0 & 0 \notin [0,2]. \end{cases}$

Another example

- Suppose $f(x) = \begin{cases} x/2 & x \in [0,2] \\ 0 & 0 \notin [0,2]. \end{cases}$
- What is $P\{X < 3/2\}$?

Another example

- Suppose $f(x) = \begin{cases} x/2 & x \in [0,2] \\ 0 & 0 \notin [0,2]. \end{cases}$
- What is $P\{X < 3/2\}$?
- What is $P\{X = 3/2\}$?

Another example

- Suppose $f(x) = \begin{cases} x/2 & x \in [0,2] \\ 0 & 0 \notin [0,2]. \end{cases}$
- What is $P\{X < 3/2\}$?
- What is $P\{X = 3/2\}$?
- What is $P\{1/2 < X < 3/2\}$?

Another example

- Suppose $f(x) = \begin{cases} x/2 & x \in [0,2] \\ 0 & 0 \notin [0,2]. \end{cases}$
- What is $P\{X < 3/2\}$?
- What is $P\{X = 3/2\}$?
- What is $P\{1/2 < X < 3/2\}$?
- What is F?

Outline

Outline

Expectations of continuous random variables

- Recall that when X was a discrete random variable, with $p(x) = P\{X = x\}$, we wrote
$$E[X] = \sum_{x:p(x)>0} p(x)x.$$

Expectations of continuous random variables

- Recall that when X was a discrete random variable, with $p(x) = P\{X = x\}$, we wrote
$$E[X] = \sum_{x:p(x)>0} p(x)x.$$
- How should we define $E[X]$ when X is a continuous random variable?

Expectations of continuous random variables

- Recall that when X was a discrete random variable, with $p(x) = P\{X = x\}$, we wrote

$$E[X] = \sum_{x:p(x)>0} p(x)x.$$

- How should we define $E[X]$ when X is a continuous random variable?
- Answer: $E[X] = \int_{-\infty}^{\infty} f(x)xdx$.

Expectations of continuous random variables

- Recall that when X was a discrete random variable, with $p(x) = P\{X = x\}$, we wrote
$$E[X] = \sum_{x:p(x)>0} p(x)x.$$
- How should we define $E[X]$ when X is a continuous random variable?
- Answer: $E[X] = \int_{-\infty}^{\infty} f(x)xdx$.
- Recall that when X was a discrete random variable, with $p(x) = P\{X = x\}$, we wrote
$$E[g(X)] = \sum_{x:p(x)>0} p(x)g(x).$$

Expectations of continuous random variables

- Recall that when X was a discrete random variable, with $p(x) = P\{X = x\}$, we wrote

$$E[X] = \sum_{x:p(x)>0} p(x)x.$$

- How should we define $E[X]$ when X is a continuous random variable?
- Answer: $E[X] = \int_{-\infty}^{\infty} f(x)xdx$.
- Recall that when X was a discrete random variable, with $p(x) = P\{X = x\}$, we wrote

$$E[g(X)] = \sum_{x:p(x)>0} p(x)g(x).$$

- What is the analog when X is a continuous random variable?

Expectations of continuous random variables

- Recall that when X was a discrete random variable, with $p(x) = P\{X = x\}$, we wrote
$$E[X] = \sum_{x:p(x)>0} p(x)x.$$
- How should we define $E[X]$ when X is a continuous random variable?
- Answer: $E[X] = \int_{-\infty}^{\infty} f(x)xdx$.
- Recall that when X was a discrete random variable, with $p(x) = P\{X = x\}$, we wrote
$$E[g(X)] = \sum_{x:p(x)>0} p(x)g(x).$$
- What is the analog when X is a continuous random variable?
- Answer: we will write $E[g(X)] = \int_{-\infty}^{\infty} f(x)g(x)dx$.

Variance of continuous random variables

- Suppose X is a continuous random variable with mean μ.

Variance of continuous random variables

- Suppose X is a continuous random variable with mean μ.
- We can write $\mathrm{Var}[X] = E[(X-\mu)^2]$, same as in the discrete case.

Variance of continuous random variables

- Suppose X is a continuous random variable with mean μ.
- We can write $\mathrm{Var}[X] = E[(X-\mu)^2]$, same as in the discrete case.
- Next, if $g = g_1 + g_2$ then $E[g(X)] = \int g_1(x)f(x)dx + \int g_2(x)f(x)dx = \int \big(g_1(x) + g_2(x)\big) f(x)dx = E[g_1(X)] + E[g_2(X)]$.

Variance of continuous random variables

- Suppose X is a continuous random variable with mean μ.
- We can write $\mathrm{Var}[X] = E[(X-\mu)^2]$, same as in the discrete case.
- Next, if $g = g_1 + g_2$ then $E[g(X)] = \int g_1(x)f(x)dx + \int g_2(x)f(x)dx = \int \big(g_1(x) + g_2(x)\big) f(x)dx = E[g_1(X)] + E[g_2(X)]$.
- Furthermore, $E[ag(X)] = aE[g(X)]$ when a is a constant.

Variance of continuous random variables

- Suppose X is a continuous random variable with mean μ.
- We can write $\mathrm{Var}[X] = E[(X-\mu)^2]$, same as in the discrete case.
- Next, if $g = g_1 + g_2$ then $E[g(X)] = \int g_1(x)f(x)dx + \int g_2(x)f(x)dx = \int \big(g_1(x) + g_2(x)\big) f(x)dx = E[g_1(X)] + E[g_2(X)]$.
- Furthermore, $E[ag(X)] = aE[g(X)]$ when a is a constant.
- Just as in the discrete case, we can expand the variance expression as $\mathrm{Var}[X] = E[X^2 - 2\mu X + \mu^2]$ and use additivity of expectation to say that $\mathrm{Var}[X] = E[X^2] - 2\mu E[X] + E[\mu^2] = E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - E[X]^2$.

Variance of continuous random variables

- Suppose X is a continuous random variable with mean μ.
- We can write $\mathrm{Var}[X] = E[(X-\mu)^2]$, same as in the discrete case.
- Next, if $g = g_1 + g_2$ then $E[g(X)] = \int g_1(x)f(x)dx + \int g_2(x)f(x)dx = \int \big(g_1(x) + g_2(x)\big)f(x)dx = E[g_1(X)] + E[g_2(X)]$.
- Furthermore, $E[ag(X)] = aE[g(X)]$ when a is a constant.
- Just as in the discrete case, we can expand the variance expression as $\mathrm{Var}[X] = E[X^2 - 2\mu X + \mu^2]$ and use additivity of expectation to say that $\mathrm{Var}[X] = E[X^2] - 2\mu E[X] + E[\mu^2] = E[X^2] - 2\mu^2 + \mu^2 = E[X^2] - E[X]^2$.
- This formula is often useful for calculations.

Examples

- Suppose that $f_X(x) = \begin{cases} 1/2 & x \in [0,2] \\ 0 & x \notin [0,2]. \end{cases}$.

Examples

- Suppose that $f_X(x) = \begin{cases} 1/2 & x \in [0,2] \\ 0 & x \notin [0,2]. \end{cases}$.
- What is $\mathrm{Var}[X]$?

Examples

- Suppose that $f_X(x) = \begin{cases} 1/2 & x \in [0,2] \\ 0 & x \notin [0,2]. \end{cases}$.
- What is $\mathrm{Var}[X]$?
- Suppose instead that $f_X(x) = \begin{cases} x/2 & x \in [0,2] \\ 0 & 0 \notin [0,2]. \end{cases}$

Examples

- Suppose that $f_X(x) = \begin{cases} 1/2 & x \in [0,2] \\ 0 & x \notin [0,2]. \end{cases}$.
- What is $\mathrm{Var}[X]$?
- Suppose instead that $f_X(x) = \begin{cases} x/2 & x \in [0,2] \\ 0 & 0 \notin [0,2]. \end{cases}$
- What is $\mathrm{Var}[X]$?

Outline

Outline

Uniform measure on $[0, 1]$

- One of the very simplest probability density functions is $f(x) = \begin{cases} 1 & x \in [0, 1] \\ 0 & 0 \notin [0, 1]. \end{cases}.$

Uniform measure on $[0, 1]$

- One of the very simplest probability density functions is $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & 0 \notin [0,1]. \end{cases}$.
- If $B \subset [0,1]$ is an interval, then $P\{X \in B\}$ is the length of that interval.

Uniform measure on $[0, 1]$

- One of the very simplest probability density functions is $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & 0 \notin [0,1]. \end{cases}$.
- If $B \subset [0, 1]$ is an interval, then $P\{X \in B\}$ is the length of that interval.
- Generally, if $B \subset [0, 1]$ then $P\{X \in B\} = \int_B 1dx = \int 1_B(x)dx$ is the "total volume" or "total length" of the set B.

Uniform measure on $[0, 1]$

- One of the very simplest probability density functions is $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & 0 \notin [0,1]. \end{cases}$.
- If $B \subset [0, 1]$ is an interval, then $P\{X \in B\}$ is the length of that interval.
- Generally, if $B \subset [0, 1]$ then $P\{X \in B\} = \int_B 1dx = \int 1_B(x)dx$ is the "total volume" or "total length" of the set B.
- What if B is the set of all rational numbers?

Uniform measure on $[0, 1]$

- One of the very simplest probability density functions is $f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & 0 \notin [0,1]. \end{cases}$.
- If $B \subset [0,1]$ is an interval, then $P\{X \in B\}$ is the length of that interval.
- Generally, if $B \subset [0,1]$ then $P\{X \in B\} = \int_B 1dx = \int 1_B(x)dx$ is the "total volume" or "total length" of the set B.
- What if B is the set of all rational numbers?
- How do we mathematically define the volume of an arbitrary set B?

Do all sets have probabilities? A famous paradox:

- Uniform probability measure on $[0, 1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0, 1)$, their probabilities should be equal.

Do all sets have probabilities? A famous paradox:

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x + r) \mod 1$.

Do all sets have probabilities? A famous paradox:

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x + r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.

Do all sets have probabilities? A famous paradox:

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x+r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.
- Call x, y "equivalent modulo rationals" if $x - y$ is rational (e.g., $x = \pi - 3$ and $y = \pi - 9/4$). An **equivalence class** is the set of points in $[0,1)$ equivalent to some given point.

Do all sets have probabilities? A famous paradox:

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x+r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.
- Call x, y "equivalent modulo rationals" if $x - y$ is rational (e.g., $x = \pi - 3$ and $y = \pi - 9/4$). An **equivalence class** is the set of points in $[0,1)$ equivalent to some given point.
- There are uncountably many of these classes.

Do all sets have probabilities? A famous paradox:

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x + r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.
- Call x, y "equivalent modulo rationals" if $x - y$ is rational (e.g., $x = \pi - 3$ and $y = \pi - 9/4$). An **equivalence class** is the set of points in $[0,1)$ equivalent to some given point.
- There are uncountably many of these classes.
- Let $A \subset [0,1)$ contain **one** point from each class. For each $x \in [0,1)$, there is **one** $a \in A$ such that $r = x - a$ is rational.

Do all sets have probabilities? A famous paradox:

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x+r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.
- Call x, y "equivalent modulo rationals" if $x - y$ is rational (e.g., $x = \pi - 3$ and $y = \pi - 9/4$). An **equivalence class** is the set of points in $[0,1)$ equivalent to some given point.
- There are uncountably many of these classes.
- Let $A \subset [0,1)$ contain **one** point from each class. For each $x \in [0,1)$, there is **one** $a \in A$ such that $r = x - a$ is rational.
- Then each x in $[0,1)$ lies in $\tau_r(A)$ for **one** rational $r \in [0,1)$.

Do all sets have probabilities? A famous paradox:

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x + r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.
- Call x, y "equivalent modulo rationals" if $x - y$ is rational (e.g., $x = \pi - 3$ and $y = \pi - 9/4$). An **equivalence class** is the set of points in $[0,1)$ equivalent to some given point.
- There are uncountably many of these classes.
- Let $A \subset [0,1)$ contain **one** point from each class. For each $x \in [0,1)$, there is **one** $a \in A$ such that $r = x - a$ is rational.
- Then each x in $[0,1)$ lies in $\tau_r(A)$ for **one** rational $r \in [0,1)$.
- Thus $[0,1) = \cup \tau_r(A)$ as r ranges over rationals in $[0,1)$.

Do all sets have probabilities? A famous paradox:

- Uniform probability measure on $[0,1)$ should satisfy **translation invariance:** If B and a horizontal translation of B are both subsets $[0,1)$, their probabilities should be equal.
- Consider **wrap-around translations** $\tau_r(x) = (x+r) \mod 1$.
- By translation invariance, $\tau_r(B)$ has same probability as B.
- Call x, y "equivalent modulo rationals" if $x - y$ is rational (e.g., $x = \pi - 3$ and $y = \pi - 9/4$). An **equivalence class** is the set of points in $[0,1)$ equivalent to some given point.
- There are uncountably many of these classes.
- Let $A \subset [0,1)$ contain **one** point from each class. For each $x \in [0,1)$, there is **one** $a \in A$ such that $r = x - a$ is rational.
- Then each x in $[0,1)$ lies in $\tau_r(A)$ for **one** rational $r \in [0,1)$.
- Thus $[0,1) = \cup \tau_r(A)$ as r ranges over rationals in $[0,1)$.
- If $P(A) = 0$, then $P(S) = \sum_r P(\tau_r(A)) = 0$. If $P(A) > 0$ then $P(S) = \sum_r P(\tau_r(A)) = \infty$. Contradicts $P(S) = 1$ axiom.

Three ways to get around this

- 1. **Re-examine axioms of mathematics:** the very *existence* of a set A with one element from each equivalence class is consequence of so-called **axiom of choice**. Removing that axiom makes paradox goes away, since one can just suppose (pretend?) these kinds of sets don't exist.

Three ways to get around this

- 1. **Re-examine axioms of mathematics:** the very *existence* of a set A with one element from each equivalence class is consequence of so-called **axiom of choice**. Removing that axiom makes paradox goes away, since one can just suppose (pretend?) these kinds of sets don't exist.
- 2. **Re-examine axioms of probability:** Replace *countable additivity* with *finite additivity*? (Look up Banach-Tarski.)

Three ways to get around this

- 1. **Re-examine axioms of mathematics:** the very *existence* of a set A with one element from each equivalence class is consequence of so-called **axiom of choice**. Removing that axiom makes paradox goes away, since one can just suppose (pretend?) these kinds of sets don't exist.
- 2. **Re-examine axioms of probability:** Replace *countable additivity* with *finite additivity*? (Look up Banach-Tarski.)
- 3. **Keep the axiom of choice and countable additivity but don't define probabilities of all sets:** Instead of defining $P(B)$ for *every* subset B of sample space, restrict attention to a family of so-called **"measurable"** sets.

Three ways to get around this

- 1. **Re-examine axioms of mathematics:** the very *existence* of a set A with one element from each equivalence class is consequence of so-called **axiom of choice**. Removing that axiom makes paradox goes away, since one can just suppose (pretend?) these kinds of sets don't exist.
- 2. **Re-examine axioms of probability:** Replace *countable additivity* with *finite additivity*? (Look up Banach-Tarski.)
- 3. **Keep the axiom of choice and countable additivity but don't define probabilities of all sets:** Instead of defining $P(B)$ for *every* subset B of sample space, restrict attention to a family of so-called **"measurable"** sets.
- Most mainstream probability and analysis takes the third approach.

Three ways to get around this

- 1. **Re-examine axioms of mathematics:** the very *existence* of a set A with one element from each equivalence class is consequence of so-called **axiom of choice**. Removing that axiom makes paradox goes away, since one can just suppose (pretend?) these kinds of sets don't exist.
- 2. **Re-examine axioms of probability:** Replace *countable additivity* with *finite additivity*? (Look up Banach-Tarski.)
- 3. **Keep the axiom of choice and countable additivity but don't define probabilities of all sets:** Instead of defining $P(B)$ for *every* subset B of sample space, restrict attention to a family of so-called **"measurable"** sets.
- Most mainstream probability and analysis takes the third approach.
- In practice, sets we care about (e.g., countable unions of points and intervals) tend to be measurable.

Perspective

- More advanced courses in probability and analysis (such as 18.125 and 18.175) spend a significant amount of time rigorously constructing a class of so-called **measurable sets** and the so-called **Lebesgue measure**, which assigns a real number (a measure) to each of these sets.

Perspective

- More advanced courses in probability and analysis (such as 18.125 and 18.175) spend a significant amount of time rigorously constructing a class of so-called **measurable sets** and the so-called **Lebesgue measure**, which assigns a real number (a measure) to each of these sets.
- These courses also replace the **Riemann integral** with the so-called **Lebesgue integral**.

Perspective

- More advanced courses in probability and analysis (such as 18.125 and 18.175) spend a significant amount of time rigorously constructing a class of so-called **measurable sets** and the so-called **Lebesgue measure**, which assigns a real number (a measure) to each of these sets.
- These courses also replace the **Riemann integral** with the so-called **Lebesgue integral**.
- We will not treat these topics any further in this course.

Perspective

- More advanced courses in probability and analysis (such as 18.125 and 18.175) spend a significant amount of time rigorously constructing a class of so-called **measurable sets** and the so-called **Lebesgue measure**, which assigns a real number (a measure) to each of these sets.
- These courses also replace the **Riemann integral** with the so-called **Lebesgue integral**.
- We will not treat these topics any further in this course.
- We usually limit our attention to probability density functions f and sets B for which the ordinary Riemann integral $\int 1_B(x)f(x)dx$ is well defined.

Perspective

- More advanced courses in probability and analysis (such as 18.125 and 18.175) spend a significant amount of time rigorously constructing a class of so-called **measurable sets** and the so-called **Lebesgue measure**, which assigns a real number (a measure) to each of these sets.
- These courses also replace the **Riemann integral** with the so-called **Lebesgue integral**.
- We will not treat these topics any further in this course.
- We usually limit our attention to probability density functions f and sets B for which the ordinary Riemann integral $\int 1_B(x)f(x)dx$ is well defined.
- Riemann integration is a mathematically rigorous theory. It's just not as robust as Lebesgue integration.