

Continuous Random Variables

- Recall: For discrete random variables, only a finite or countably infinite number of possible values with positive probability (>0).

➤ Often, there is interest in random variables that can take (at least theoretically) on an uncountable number of possible values,

e.g.,

Note: observed data, always discrete

$(0, 800)$ ← the weight of a randomly selected person in a population,

$(0, \infty)$ ← the length of time that a randomly selected light bulb works,

$(-a, a)$ ← the error in experimentally measuring the speed of light.

for some a ,
or $(-\infty, \infty)$

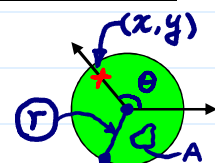
➤ Example (Uniform Spinner, LNp.3-6, 3-18):

a probability space
 $(\Omega, \mathcal{F}, P)$

■ $\Omega = (-\pi, \pi]$ or $\Omega^* = \{(x, y) \mid x^2 + y^2 \leq r^2\}$

■ For $(a, b] \subset \Omega$, $P((a, b]) = (b-a)/(2\pi)$ or $P(A)$

$= \text{area}(A)/\pi r^2$
for $A \subset \Omega^*$



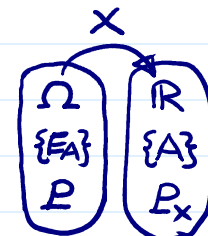
$P(\{\omega\}) = 0$
for $\forall \omega \in \Omega$

■ Consider the random variables:

$X: \Omega \rightarrow \mathbb{R}$, and $X(\omega) = \omega$ for $\omega \in \Omega$,

$\Omega^* \rightarrow \mathbb{R}$
 $(x, y) \rightarrow \theta$

Range of X : $(-\pi, \pi]$ ← uncountable set



$Y: \Omega \rightarrow \mathbb{R}$, and $Y(\omega) = \tan(\omega)$ for $\omega \in \Omega$.

Range of Y : $(-\infty, \infty)$ ← uncountable set

$\Omega \rightarrow \mathbb{R}$ or $\Omega^* \rightarrow \mathbb{R}$

Then, X and Y are random variables that takes on an uncountable number of possible values. ← defined by the "range of r.v."

Recall.
LNp.3-18

■ Some properties about the distribution of X (or Y), i.e. P_X (or P_Y)

■ $P_X(\{X = x\}) = P(\{x\}) = 0$, for any $x \in \mathbb{R}$.

⇒ Probability for X to take any single value is zero

■ But, for $-\pi \leq a < b \leq \pi$,

$P_X(\{X \in (a, b]\}) = P((a, b]) = (b-a)/(2\pi) > 0$.

⇒ Positive probability (>0) is assigned to any $(a, b]$

Why it happen?
Note: uncountable sum.

■ Q: Can we still define a probability mass function for X ?

■ Q: If not, what can play a similar role like pmf for X ?

discrete pmf countable $\sum$	continuous pdf uncountable $\int$
-------------------------------------	---

Recall. Find area under a curve

by integration

(uncountable sum)

$f(x) \geq 0$

$\text{area}((a, b)) = \int_a^b f(x) dx$

$f(x) > 0$, but
 $\text{area}(\{c\}) = \int_c^c f(x) dx = 0$



No. $\because$
 $\text{pmf}(x) = P_X(X=x)$

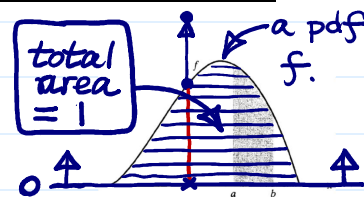
• Probability Density Function and Continuous Random Variable

p. 6-3

↳ it plays a similar role like pmf for discrete r.v

➤ Definition. A function $f: \mathbb{R} \rightarrow \mathbb{R}$ is called a probability density function (pdf) if

density: 密度, mass = density × volume



LNp.5-6,
(i)(ii)(iii)

C.F.

1. $f(x) \geq 0$, for all $x \in (-\infty, \infty)$, and

2. $\int_{-\infty}^{\infty} f(x) dx = 1$.

f : not necessary to be a continuous function

➤ Definition: A random variable X is called continuous if there exists a pdf f such that for any set B of real numbers

$P(X \in B) = \sum_{X \in B \cap \mathbb{N}} f(x)$

LNp.5-6,
(iv)

$P_X(\{X \in B\}) = \int_B f(x) dx$.

can be regarded as a small prob.

$f(x)$: not probability

not unique

area = probability

C.F.

area

$f(x)$

$x - \frac{dx}{2}$ $x + \frac{dx}{2}$

For example, $P_X(a \leq X \leq b) = \int_a^b f(x) dx$.

$B = [a, b]$

C.F.

Theorem in LNp.5-7

Theorem. If f is a pdf, then there must exist a continuous random variable with pdf f .

red area = $f(x) dx \approx P_X(x - \frac{dx}{2} < X < x + \frac{dx}{2})$

Sketch of proof. Let $F(x) = \int_{-\infty}^x f(t) dt$, then show

that $F(x)$ is a cdf (exercise, check LNp.5-9-10, (2)(3)(4)).

Then, by the theorem in LNp.5-11, there exist a r.v. X s.t. $P_X((a, b]) = F(b) - F(a) = \int_a^b f(x) dx$.

Why? $\because F_X(x)$ is defined as $P_X(X \in (-\infty, x])$

Some properties

Some properties $\because F_X(x) = P(X \in (-\infty, x]) = P(X \in (-\infty, x)) = F_X(x-)$

The cdf of X is a continuous function, i.e. no jumps.

$P_X(\{X = x\}) = \int_x^x f(y) dy = 0$ for any $x \in \mathbb{R}$

$\int_a^b f(x) dx$

It does not matter whether the intervals are open or close, i.e.,

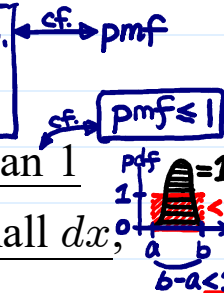
$P(X \in [a, b]) = P(X \in (a, b]) = P(X \in [a, b)) = P(X \in (a, b))$.

Q: Is the pdf of a continuous r.v. X unique? Hint. Consider $f^*(x) = f(x)$ except at countably many x 's

It is important to remember that the value of a pdf $f(x)$ is NOT a probability itself

area = prob. but $f(x)$ is not prob.

It is quite possible for a pdf to have value greater than 1



Q: How to interpret the value of a pdf $f(x)$? For small dx ,

$$P\left(x - \frac{dx}{2} \leq X \leq x + \frac{dx}{2}\right) = \int_{x - \frac{dx}{2}}^{x + \frac{dx}{2}} f(y) dy \approx f(x) \cdot dx.$$

Ans in LNp.5-5

$\Rightarrow f(x) dx$ is a measure of how likely it is that X will be near x

We can characterize the distribution of a continuous random variable in terms of its

Note cdf is defined for any r.v.'s (LNp.5-8)

1. Probability Density Function (pdf)

2. (Cumulative) Distribution Function (cdf)

3. Moment Generating Function (mgf, Chapter 7)

i.e. if $f(a) > f(b)$, then X is more likely to take value near a than value near b , e.g. $\frac{f(a)}{f(b)} = 2$ (meaning?)