Lesson 99 - Continuous Random Variables

HL Math - Santowski

Consider the following table of heights of palm trees, measured in mm, divided into intervals of 1000’s of mm,

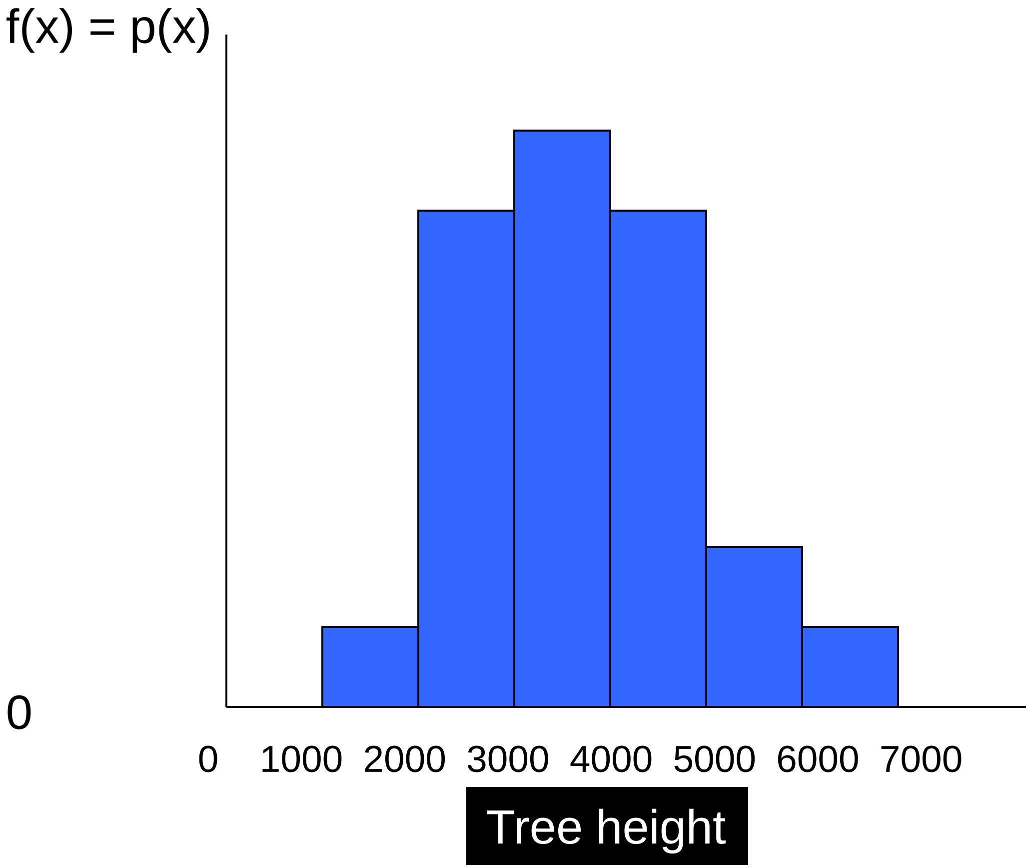
interval	Frequency
(0,1000]	0
(1000,2000]	50
(2000,3000]	250
(3000,4000]	300
(4000,5000]	250
(5000,6000]	100
(6000,7000]	50

and the relative frequency of each interval.

interval	relative freq.
(0,1000]	0
(1000,2000]	0.05
(2000,3000]	0.25
(3000,4000]	0.30
(4000,5000]	0.25
(5000,6000]	0.10
(6000,7000]	0.05
	1.00

Graph

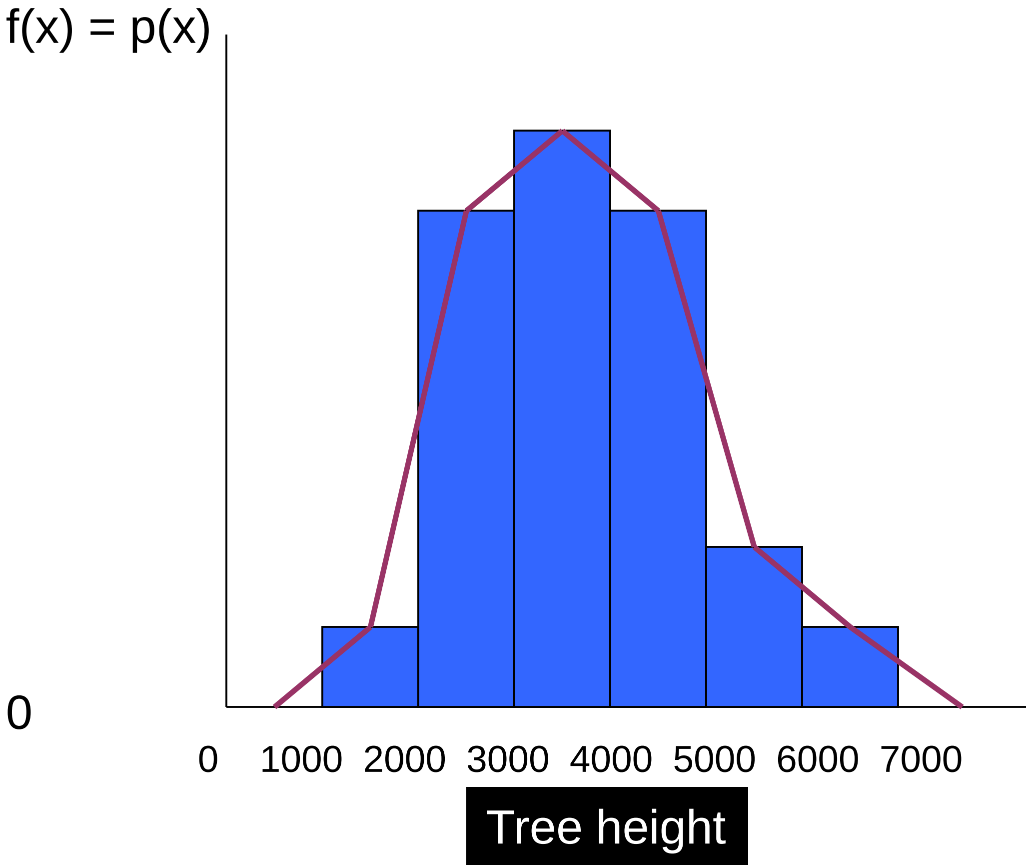
interval	relative freq.
(0,1000]	0
(1000,2000]	0.05
(2000,3000]	0.25
(3000,4000]	0.30
(4000,5000]	0.25
(5000,6000]	0.10
(6000,7000]	0.05



The area of each bar is the frequency of the category.

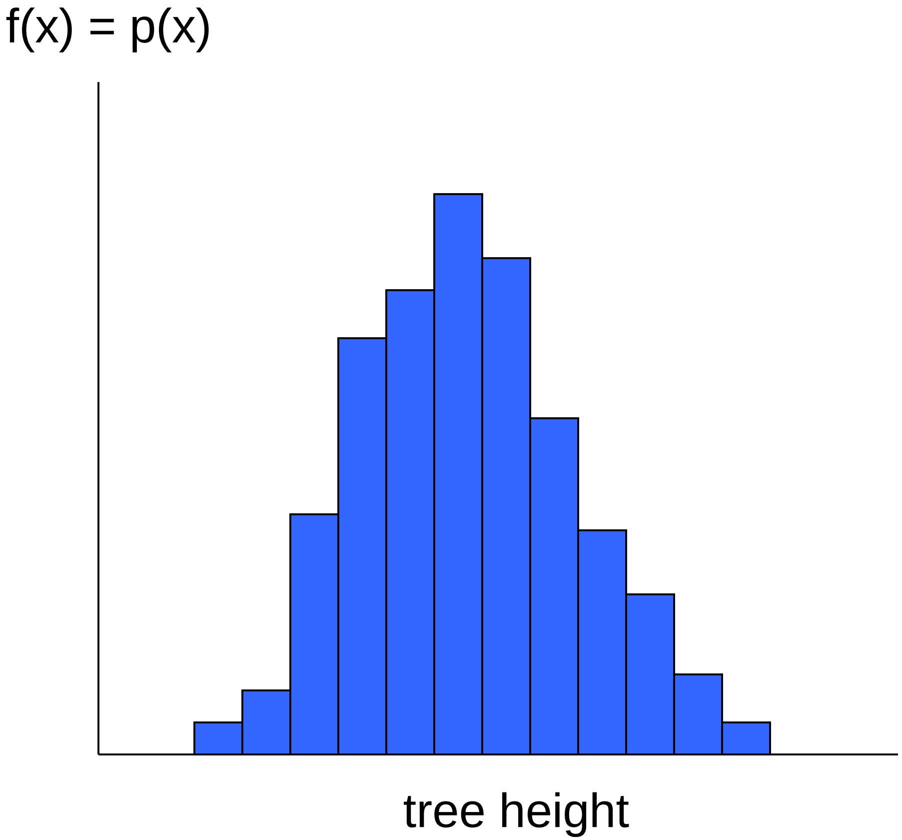
Graph

interval	relative freq.
(0,1000]	0
(1000,2000]	0.05
(2000,3000]	0.25
(3000,4000]	0.30
(4000,5000]	0.25
(5000,6000]	0.10
(6000,7000]	0.05



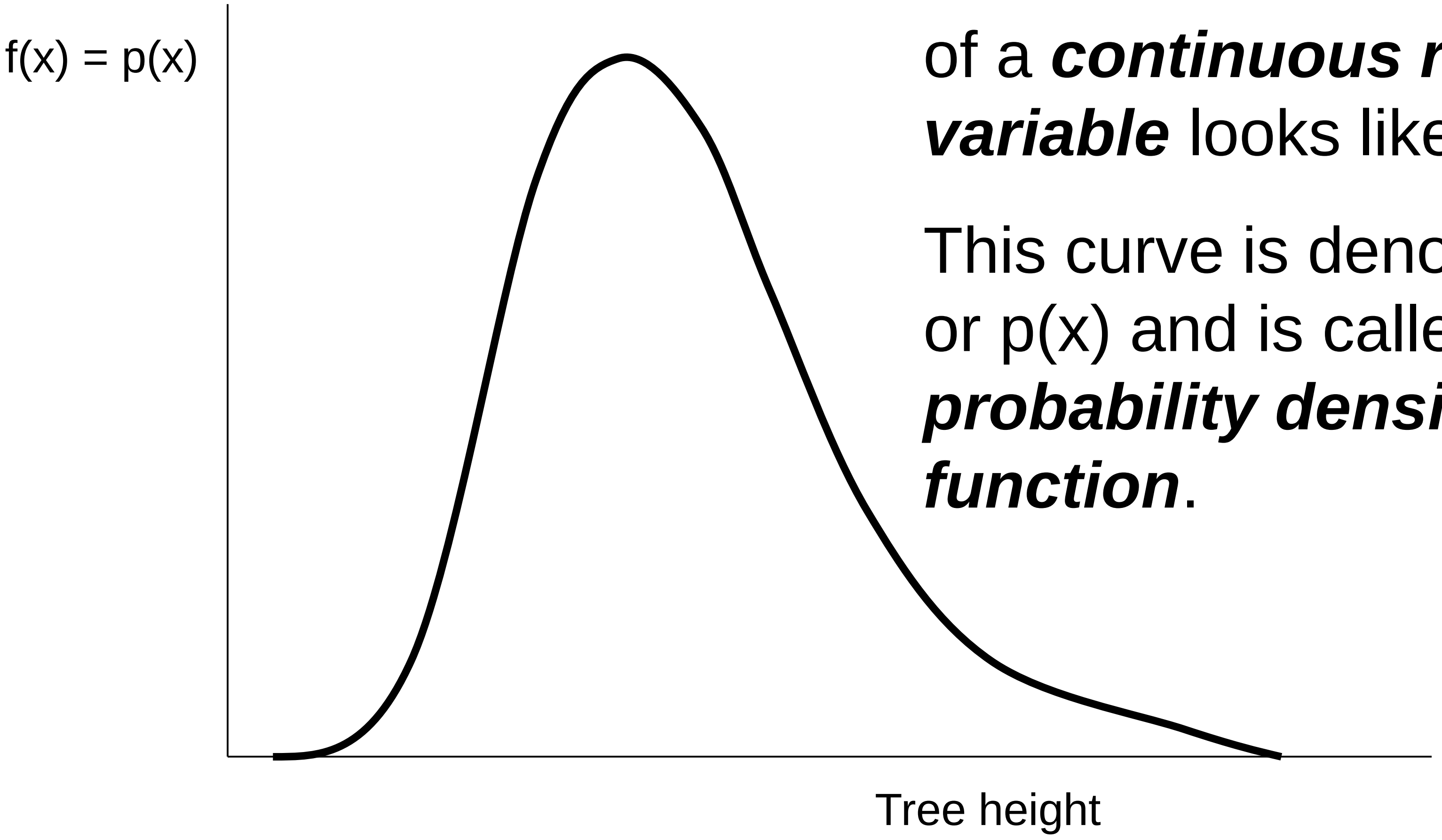
Here is the frequency polygon.

If we make the intervals 500 mm instead of 1000 mm, the graph would probably look something like this:



The height of the bars increases and decreases more gradually.

If we made the intervals infinitesimally small, the bars and the frequency polygon would become smooth, looking something like this:



This what the distribution of a ***continuous random variable*** looks like.

This curve is denoted f(x) or p(x) and is called the ***probability density function***.

pmf versus pdf

For a discrete random variable, we had a probability mass function (pmf).

The pmf looked like a bunch of spikes, and probabilities were represented by the heights of the spikes.

For a continuous random variable, we have a probability density function (pdf).

The pdf looks like a curve.

Continuous Random Variables

- A random variable was a numerical value associated with the outcome of an experiment.
  - Finite discrete random variables were ones in which the values were countable whole numbered values
- A ***continuous random variable*** is a random variable that can assume any value in some interval of numbers, and are thus NOT countable.
  - Examples:
    - The time that a train arrives at a specified stop
    - The lifetime of a transistor
    - A randomly selected number between 0 and 1
    - Let R be a future value of a weekly ratio of closing prices for IBM stock
    - Let W be the exact weight of a randomly selected student

Difference between discrete and continuous random variables

- 1. They are used to describe different types of quantities.
- 2. We use **distinct values** for discrete random variables but **continuous real numbers** for continuous random variables.
- 3. Numbers between the values of discrete random variable makes no sense, for example, P(0)=0.5, P(1)=0.5, then P(1.5) has no meaning at all. But that is not true for continuous random variables.

Difference between discrete and continuous random variables

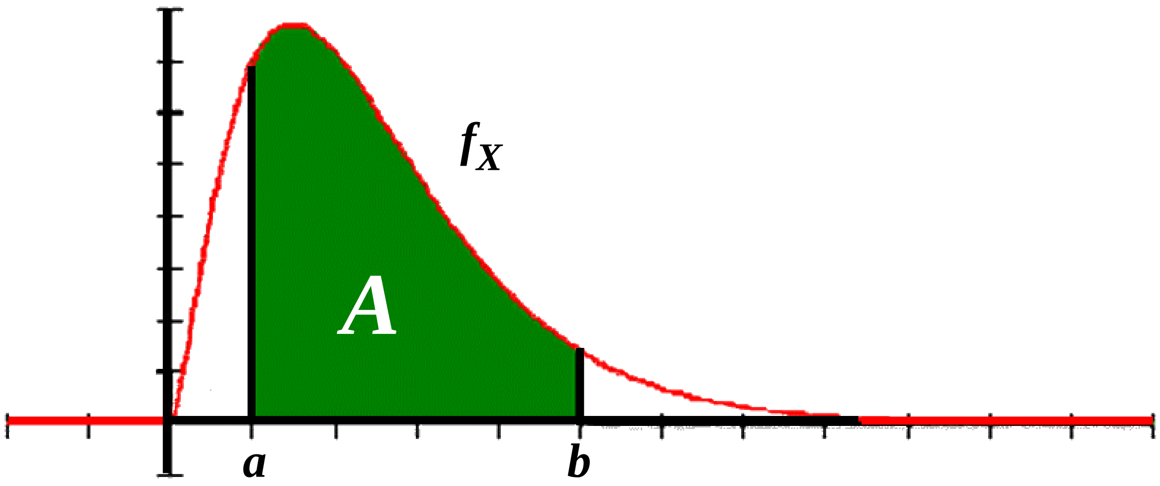
- Both discrete and continuous random variables have sample space.
- For discrete r.v., there may be finite or infinite number of sample points in the sample space.
- For continuous r.v., there are **always infinitely many** sample points in the sample space.

Difference between discrete and continuous random variables

- *** For discrete r.v., given the pmf, we can find the probability of each sample point in the sample space.
- *** But for continuous r.v., we **DO NOT** consider the probability of each sample point in the sample space because it is defined to be **ZERO!**

Continuous Random Variables

- A random variable is said to be continuous if there is a function $f_X(x)$ with the following properties:
  - Domain: all real numbers
  - Range: $f_X(x) \geq 0$
  - The area under the entire curve is 1
- Such a function $f_X(x)$ is called the ***probability density function*** (abbreviated p.d.f.)
- The fact that the total area under the curve $f_X(x)$ is 1 for all X values of the random variable tells us that all probabilities are expressed in terms of the area under the curve of this function.
  - Example: If X are values on the interval from [a,b], then the $P(a \leq X \leq b)$ = area under the graph of $f_X(x)$ over the interval [a,b]



Continuous Random Variables

- Because all probabilities for a continuous random variable are described in terms of the area under the p.d.f. function, the $P(X=x) = 0$.
  - Why: the area of the p.d.f. for a single value is zero because the width of the interval is zero!
  - That is, ***for any <u>continuous</u> random variable, X, P(X = a) = 0 for every number a.*** This DOES NOT imply that X cannot take on the value a, it simply means that the probability of that event is 0.

CRV - Probability Distribution

Let X be a continuous rv. Then a *probability distribution or probability density function (pdf)* of X is a function $f(x)$ such that for any two numbers a and b,

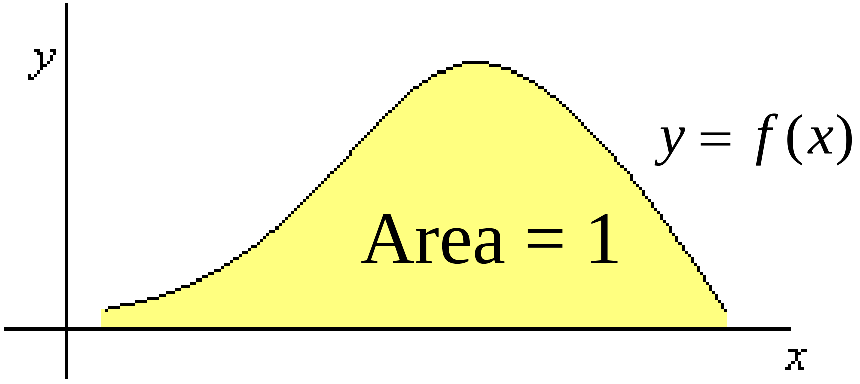
$$P\left(a \le X \le b\right) = \int_a^b f(x)dx$$

The graph of f is the *density curve*.

CRV - Probability Density Function

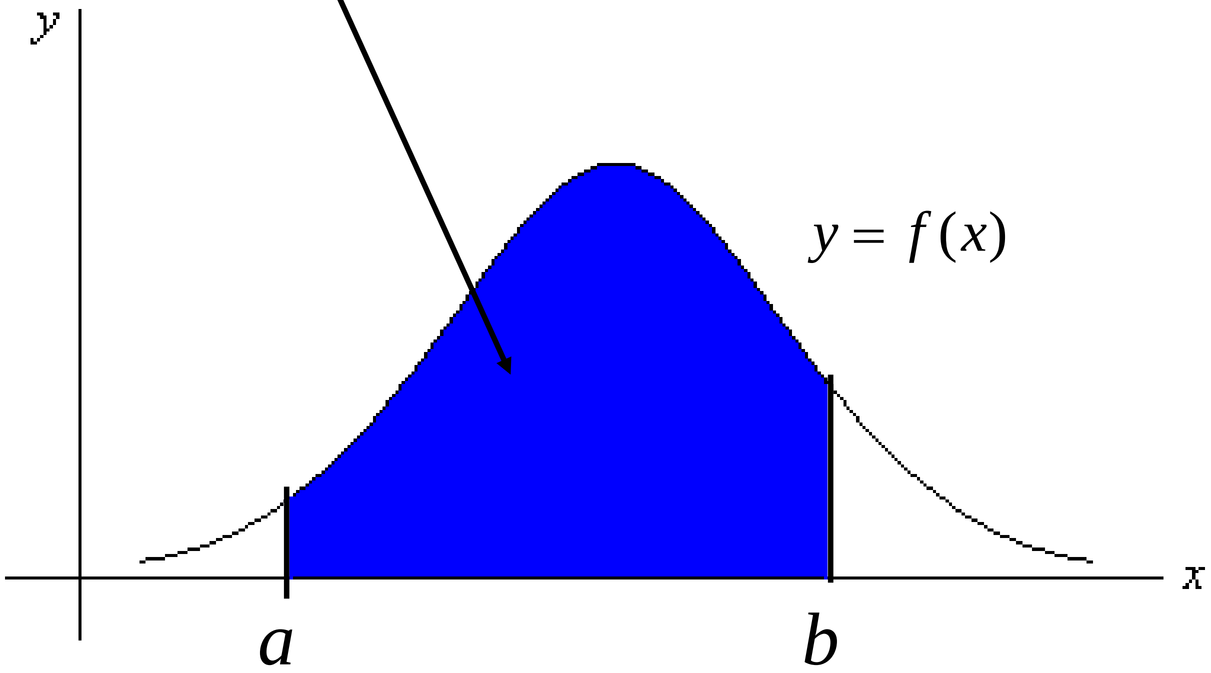
For $f(x)$ to be a pdf

1. $f(x) > 0$ for all values of x.
2. The area of the region between the graph of f and the x – axis is equal to 1.



CRV - Probability Density Function

$P(a \le X \le b)$ is given by the area of the shaded region.



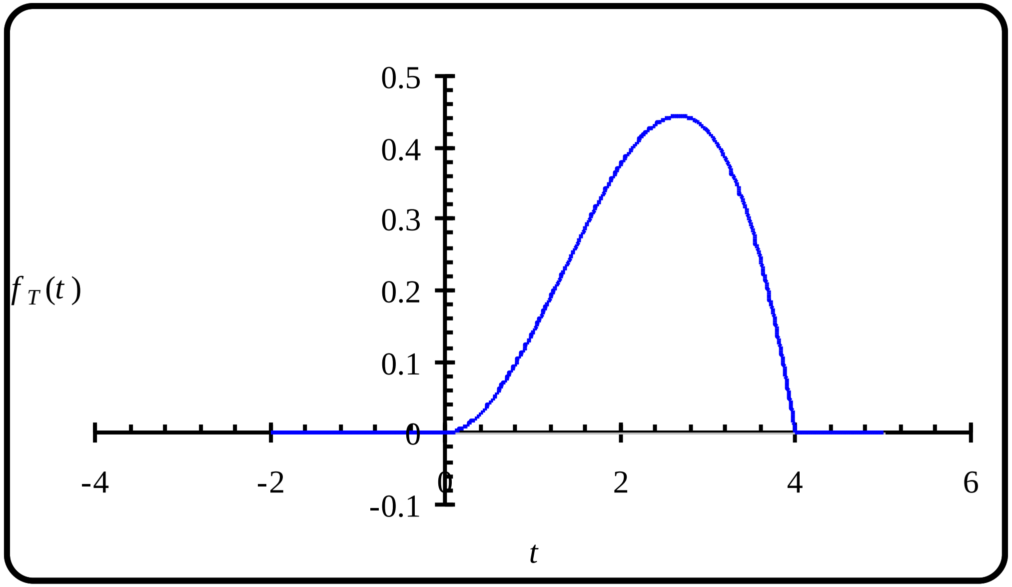
Example

- The *p.d.f*. of *T*, the weekly CPU time (in hours) used by an accounting firm, is given below.

$$f_T(t) = \begin{cases} 0 & \text{if } t < 0 \\ \frac{3}{64} t^2 (4-t) & \text{if } 0 \le t \le 4 \\ 1 & \text{if } t > 4 \end{cases}$$

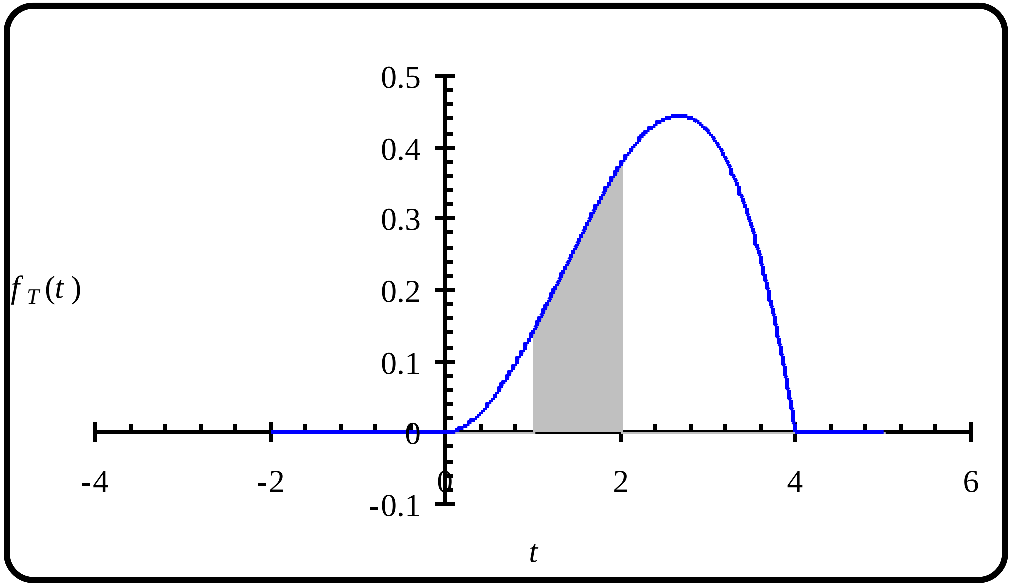
Example (cont)

- The graph of the p.d.f. is given below:



Example (cont)

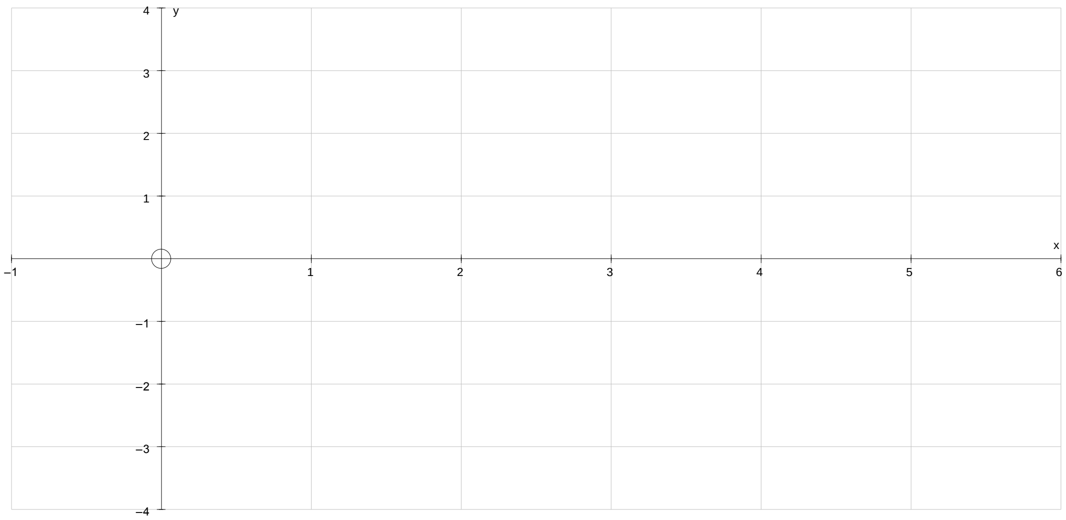
- $P(1 \le T \le 2)$ is equal to the area between the graph of and the *t*-axis over the interval.



Example

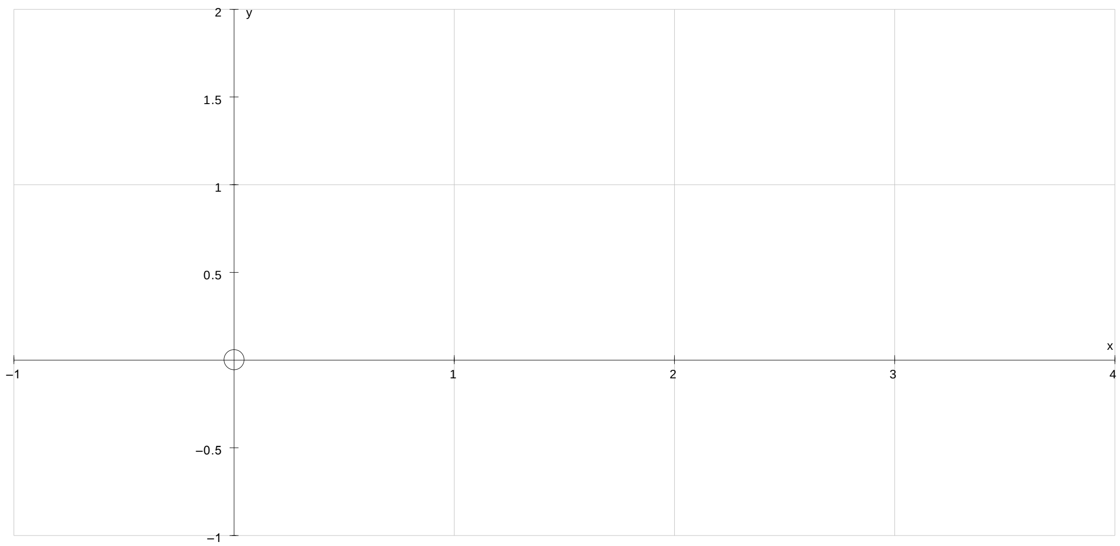
- Sketch the following functions and decide whether they are valid pdfs

$$f\left(x\right)=\begin{cases} x-2 & 0<x<5 \\ 0 & \text{otherwise} \end{cases}$$



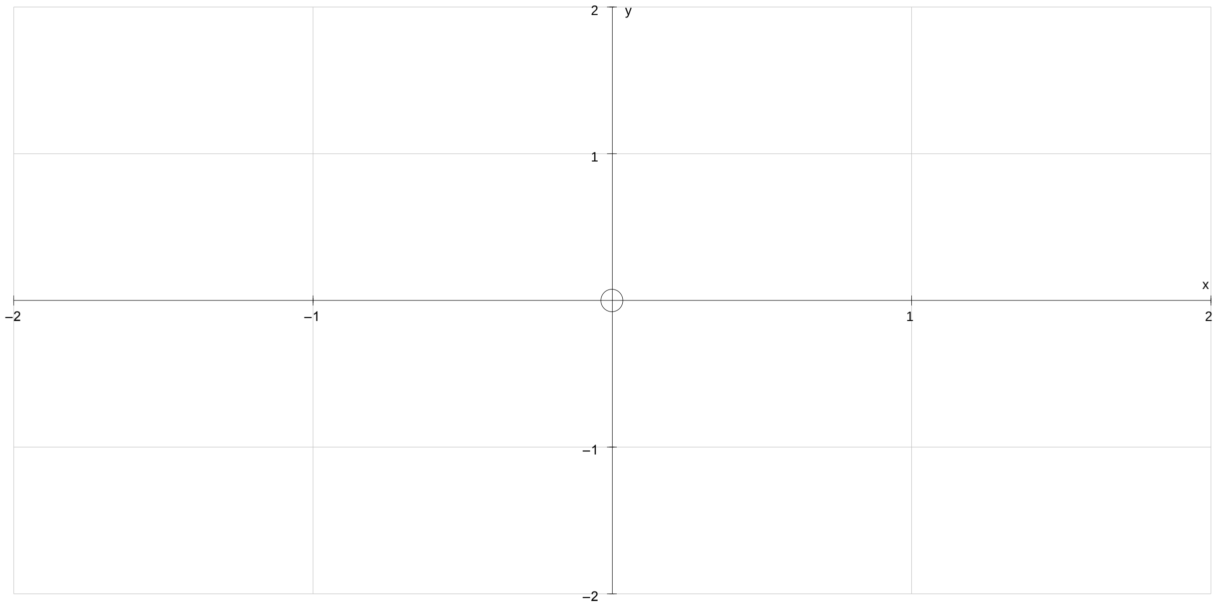
Sketch the following function and decide whether it is a valid pdf

$$f(x) = \begin{cases} \frac{1}{3}x & 0 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

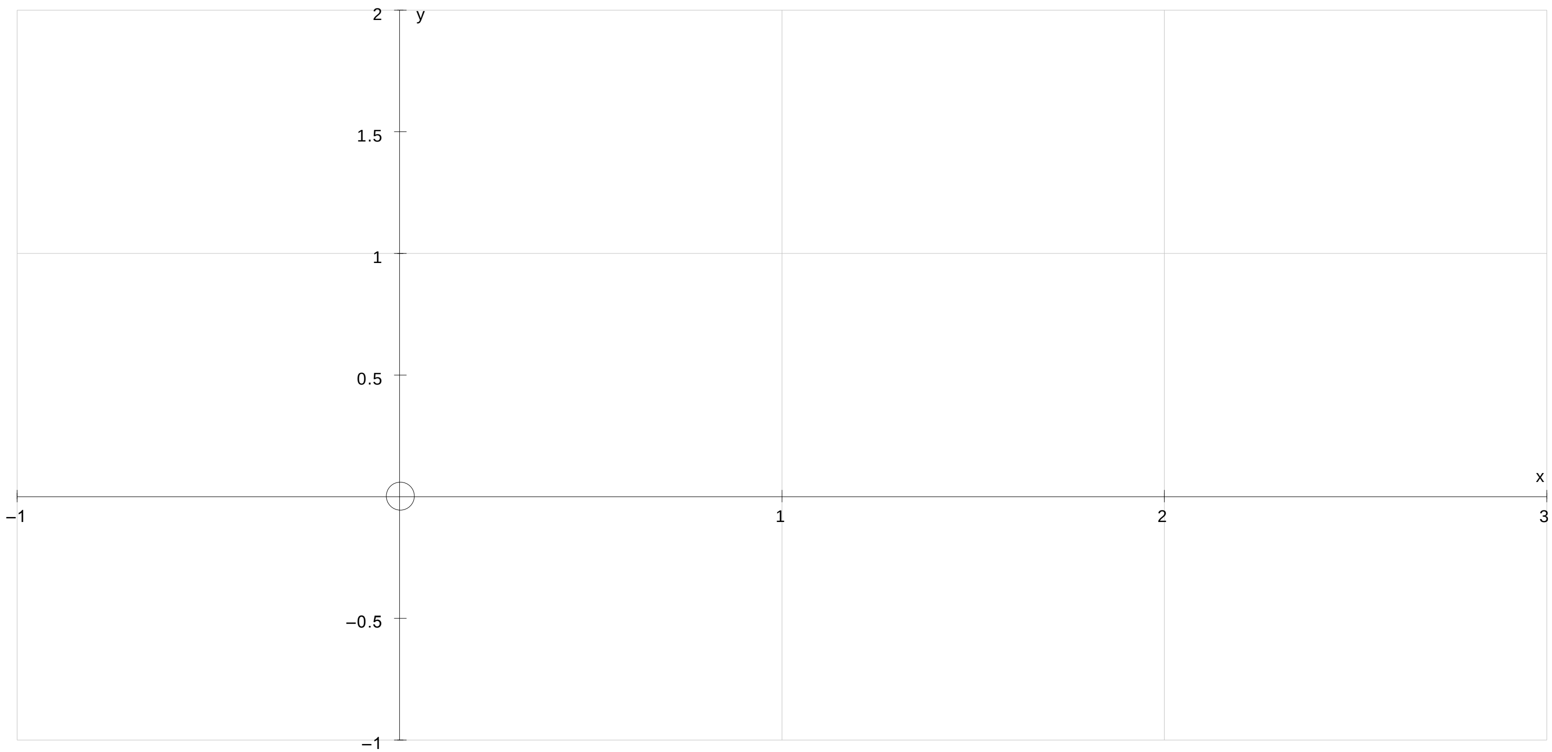


Sketch the following function and decide whether it is a valid pdf

$$f(x)=\begin{cases}\frac{3}{2}x^2 & -1<x<1\\ 0 & \text{otherwise}\end{cases}$$



$$f(x)=\begin{cases} x & 0\le x\le 1 \\ 2-x & 1\le x\le 2 \\ 0 & \text{otherwise} \end{cases}$$



Given that the following is a pdf, find the value of k

(a)

$$f(x) = \begin{cases} kx(4-x) & 2 < x < 4 \\ 0 & \text{otherwise} \end{cases}$$

(b)

$$f(x) = \begin{cases} k & 0 < x < 2 \\ k(2x-3) & 2 \le x \le 3 \\ 0 & \text{otherwise} \end{cases}$$

Example

(a)

$$f(x)=\begin{cases}\frac{1}{2}(x-3) & 3<x<5\\ 0 & \text{otherwise}\end{cases}$$

Find $P(x<4)$

$$f(x)=\begin{cases}\frac{1}{8}x & 0<x<4\\ 0 & \text{otherwise}\end{cases}$$

Find $P(1<x<2)$

Example 1 In actuarial science, one of the models used for describing mortality is

$$f(x) = \begin{cases} Cx^2(100-x)^2 & 0 \le x \le 100 \\ 0 & \text{otherwise} \end{cases},$$

where x denotes the age at which a person dies.

(a) Find the value of C.

(b) Let A be the event "Person lives past 60." Find $P(A)$.

1: Suppose X is a random variable with density function

$$f(x) = \begin{cases} 2x & \text{if } 0 < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

a) Find $P(X \leq 1/2)$.
b) Find $P(X \geq 3/4)$.
c) Find $P(X \geq 2)$.
d) Find $E[X]$.
e) Find the standard deviation of X.

2: Suppose X is a random variable with density function

$$f(x) = \begin{cases} cx^2 & \text{if } 0 < x < 2 \\ 0 & \text{otherwise} \end{cases}$$

for some positive number c. What is c?

3: Suppose the number of years that a television set lasts has density

$$f(x) = \begin{cases} 18x^{-3} & \text{if } x \geq 3 \\ 0 & \text{otherwise.} \end{cases}$$

a) Find the probability that the television set lasts between 5 and 6 years.
b) Find the probability that the television set lasts at least 4 years.
c) Find the probability that the television set lasts less than 2.5 years.
d) Find the probability that the television set lasts exactly 3.76 years.
e) Find the expected value of the number of years that the television set lasts.

4: Suppose that the number of hours that it takes for a student to finish an exam has density

$$f(x) = \begin{cases} \frac{2}{5}(x+1) & \text{if } 1 < x < 2 \\ 0 & \text{otherwise.} \end{cases}$$

a) Find the probability that the student finishes the exam in less than 1.5 hours.
b) Find the mean and standard deviation of the number of hours it takes to finish the exam.

Example I

- A continuous random variable X has the pdf $f(x)=c(x-1)(2-x)$ over the interval [1, 2] and 0 elsewhere.
  - (a) What value of c makes f(x) a valid pdf for X?
  - (b) What is $P(x>1.5)$?

Working with CRV Distributions

- Median ➔ $$\frac{1}{2} = \int_{-\infty}^{m} f(x)dx$$
- Mode ➔ max. point(s) of f(x)
- Mean (expected value) ➔ $$\int_{-\infty}^{\infty} x \cdot f(x)dx$$
- Variance ➔ $E(x^2)-[E(x)]^2$ or $$\int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)dx$$

Ex. 1: Mode from Graphs

Ex. 2: Mode from Eqn

- Determine the mode of the CRV distribution defined by:

$$f(x) = \begin{cases} \dfrac{3}{4}x^2(2-x) & 0 < x \leq 2 \\ 0 & \textbf{otherwise} \end{cases}$$

Ex. 3: Mean & Variance

- Determine the mean and variance of the CRV distribution defined by

$$f(x) = \begin{cases} \frac{3}{4}(1-x)(x-3) & 1 \le x \le 3 \\ 0 & \textbf{otherwise} \end{cases}$$

Ex. 4: Median

- Determine the median and lower quartile of the CRV distribution defined by

$$f(x)=\begin{cases}\dfrac{1}{12}(5-2x) & -1\le x\le 2\\ 0 & \textbf{otherwise}\end{cases}$$

Cumulative Distribution Functions (CDFs)

- Properties:
  - (1) $0 \leq F(x) \leq 1$
  - (2) $F(-\infty) = 0$ and then $F(+\infty) = 1$
  - (3) if $x_1 \leq x_2$, then $F(x_1) \leq F(x_2)$
  - (4) $P(x_1 < x \leq x_2) = F(x_2) - F(x_1)$
  - (5) $f(x) = d/dx\, F(x)$

Ex. 5: Given cdf, work with it ….

- A cumulative distribution function is given by:

$$F(x) = \begin{cases} 0 & \textbf{for} \;\; x < 0 \\ \dfrac{x}{2} & \textbf{for} \;\; 0 \le x < \dfrac{1}{2} \\ x^2 & \textbf{for} \;\; \dfrac{1}{2} \le x < 1 \\ 1 & \textbf{for} \;\; x > 1 \end{cases}$$

  - (a) Sketch it
  - (b) find P(x ≤ ¾)
  - (c) find P(½ < x ≤ ¾)
  - (d) find P(x > ¼)

Ex. 6: Given cdf, create pdf

Use the cumulative distribution function

$$F(x) = \begin{cases} 0 & \textbf{for} \;\; x < 0 \\ \dfrac{x}{2} & \textbf{for} \;\; 0 \le x < \dfrac{1}{2} \\ x^2 & \textbf{for} \;\; \dfrac{1}{2} \le x < 1 \\ 1 & \textbf{for} \;\; x > 1 \end{cases}$$

to determine the equation and sketch of the pdf

Ex. 7:Given pdf, draw cdf

Given the probability density function of

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \textbf{otherwise} \end{cases}$$

Draw the corresponding cumulative probability function