Chapter 4. Continuous Random Variables and Probability Distributions

Continuous Random Variables and Probability Distributions

- **4.1 Continuous Random Variables and Probability Density Functions**
- **4.2 Cumulative Distribution Functions and Expected Values**
- **4.3 The Normal Distribution**
- **4.4 The Gamma Distribution and Its Relatives**
- **4.5 Other Continuous Distributions**
- **4.6 Probability Plots**

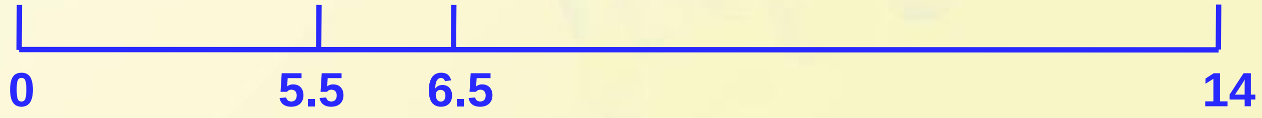
- **Continuous Random Variables**

A random variable X is said to be continuous if its set of possible values is an entire interval of numbers – that is , if for some A<B, any number x between A and B is possible

- **Example 4.2**

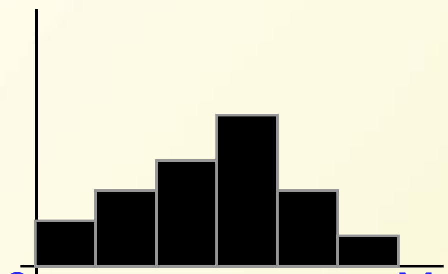
If a chemical compound is randomly selected and its PH X is determined, then X is a continuous rv because any PH value between 0 and 14 is possible.

If more is know about the compound selected for analysis, then the set of possible values might be a subinterval of [0, 14], such as $5.5 \le x \le 6.5$, but X would still be continuous.



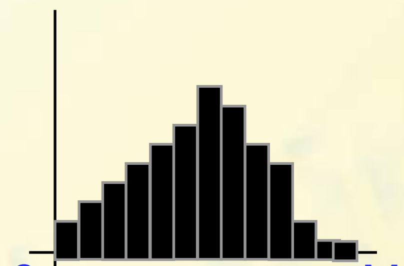
Probability Distribution for Continuous Variables

Suppose the variable X of interest is the depth of a lake at a randomly chosen point on the surface. Let M be the maximum depth, so that any number in the interval [0,M] is a possible value of X.



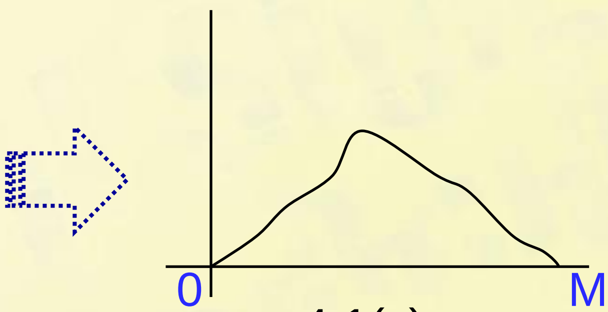
4.1(a)

Measured by meter



4.1(b)

Measured by centimeter



4.1(c)

A limit of a sequence of discrete histogram

Discrete Cases

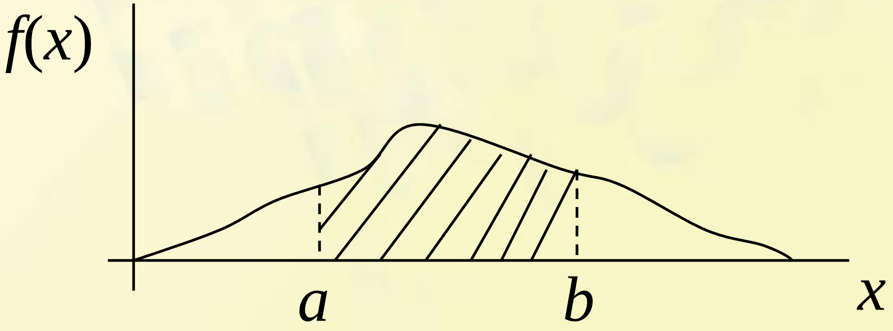
Continuous Case

Probability Distribution

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

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

The probability that X takes on a value in the interval [a,b] is the area under the graph of the density function as follows.

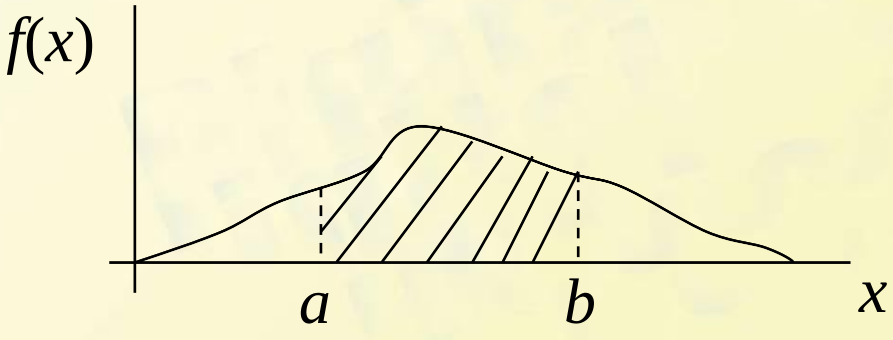


- A legitimate pdf should satisfy

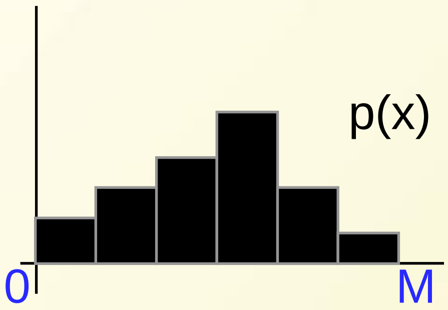
1. $f(x) \geq 0$ for all x

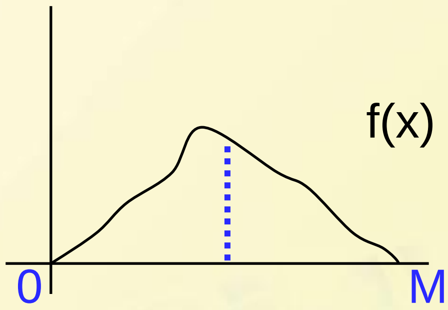
2. $\int_{-\infty}^{\infty} f(x)\mathrm{d}x =$ area under the entire graph of $f(x)$

$= 1$



- pmf (Discrete) vs. pdf (Continuous)





P(X=c) = p(c)

P(X=c) = f(c) ?

$$P(X=c)=\int_{c}^{c} f(x)dx=0$$

- **Proposition**

  If X is a continuous rv, then for any number c, $P(X=c)=0$. Furthermore, for any two numbers a and b with $a<b$,

$$
\begin{aligned}
P(a \le X \le b) &= P(a < X \le b) \\
&= P(a \le X < b) \\
&= P(a < X < b)
\end{aligned}
$$

Impossible event :the event contain no simple element

$P(A)=0$ ➔ A is an impossible event ?

Uniform Distribution

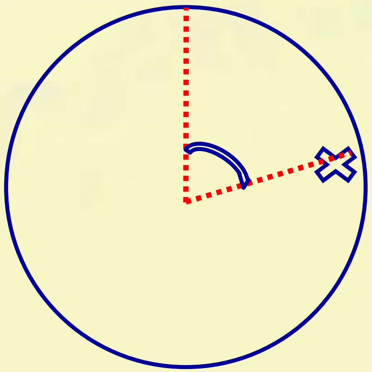
A continuous rv X is said to have a uniform distribution on the interval [A, B] if the pdf of X is

$$f(x; A, B) = \begin{cases} \dfrac{1}{B - A} & A \le x \le B \\ 0 & \text{otherwise} \end{cases}$$

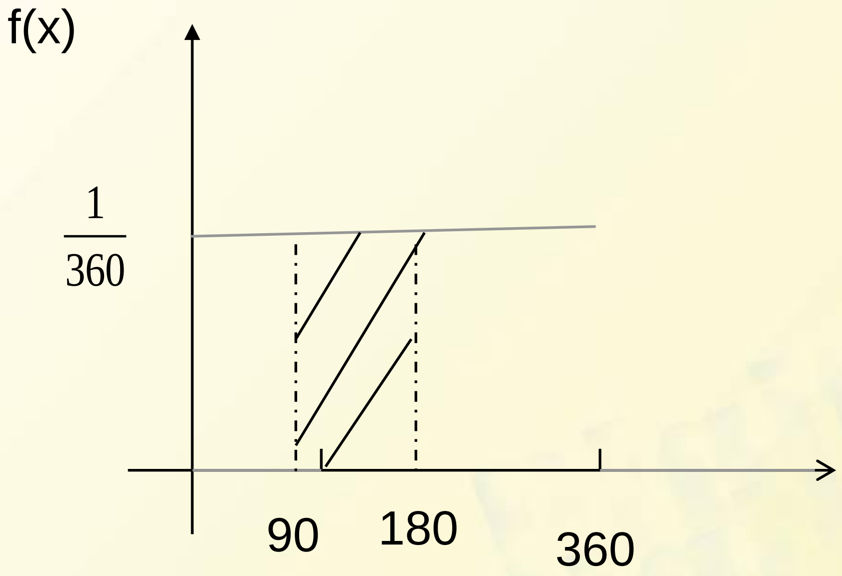
Example 4.4

The direction of an imperfection with respect to a reference line on a circular object such as a tire, brake rotor, or flywheel is, in general, subject to uncertainty. Consider the reference line connecting the valve stem on a tire to the center point, and let X be the angle measured clockwise to the location of an imperfection, One possible pdf for X is

$$f(x) = \begin{cases} \dfrac{1}{360} & 0 \le x \le 360 \\ 0 & \text{otherwise} \end{cases}$$



- **Example 4.4 (Cont')**



$$P(90 \le X \le 180) = \int_{90}^{180} \frac{1}{360}\, dx = 0.25$$

- **Example 4.5**

"Time headway" in traffic flow is the **elapsed time** between the time that one car finishes passing a fixed point and the instant that the next car begins to pass that point.

Let X = the time headway for two randomly chosen consecutive cars on a freeway during a period of heavy flow.

The following pdf of X is essentially the one suggested in "The Statistical Properties of Freeway Traffic".

$$f(x) = \begin{cases} 0.15e^{-0.15(x-0.5)} & x \geq 0.5 \\ 0 & \text{otherwise} \end{cases} \qquad \textbf{(1)}$$

(A) The formula (1) satisfy the pdf condition?

(B) The probability that headway time is at most 5 sec is ?

- **Solution:**

$$f(x)=\begin{cases}0.15\mathrm{e}^{-0.15(x-0.5)} & x\geq 0.5\\ 0 & \text{otherwise}\end{cases}$$

1. f(x) ≥ 0;

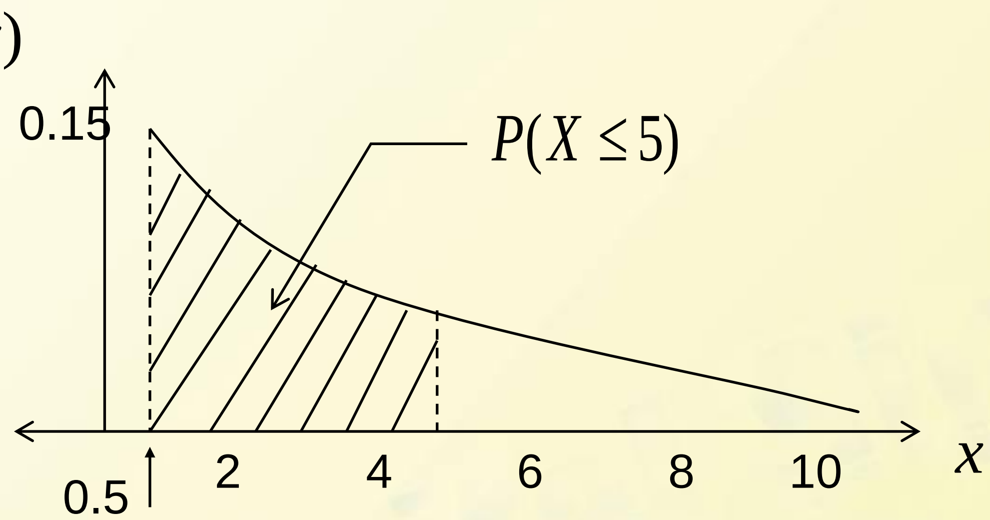
2. To show $\int_{-\infty}^{\infty} f(x) \geq 0 dx = 1$, we use the result $\int_{a}^{\infty} e^{-kx}dx=\frac{1}{k}e^{-ka}$

$$\int_{-\infty}^{\infty} f(x)dx=\int_{0.5}^{\infty}0.15e^{-0.15(x-5)}dx=0.15e^{0.075}\int_{0.5}^{\infty}e^{-0.15x}dx$$

$$=0.15e^{0.075}\cdot\frac{1}{0.15}e^{-(0.15)(0.5)}=1$$

Thus, the formula (1) satisfy the pdf condition.

- **Example 4.5 (Cont')**



$$P(X \leq 5) = \int_{-\infty}^{5} f(x)dx = \int_{0.5}^{5} .15e^{-0.15(x-5)}dx = 0.15e^{0.075}\int_{0.5}^{5} e^{-0.15x}dx$$

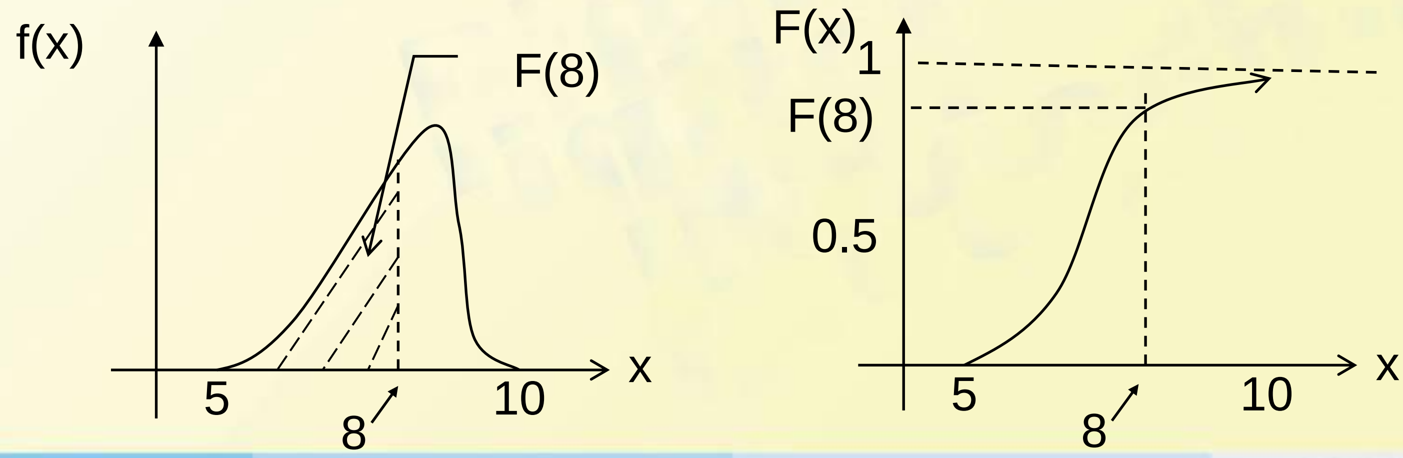
$$= 0.15e^{0.075} \cdot \left(-\frac{1}{0.15}e^{-0.15x}\right)\Bigg|_{0.5}^{5} = 0.491 = P(X < 5)$$

Cumulative Distribution Function

The cumulative distribution function F(x) for a continuous rv X is defined for every number x by

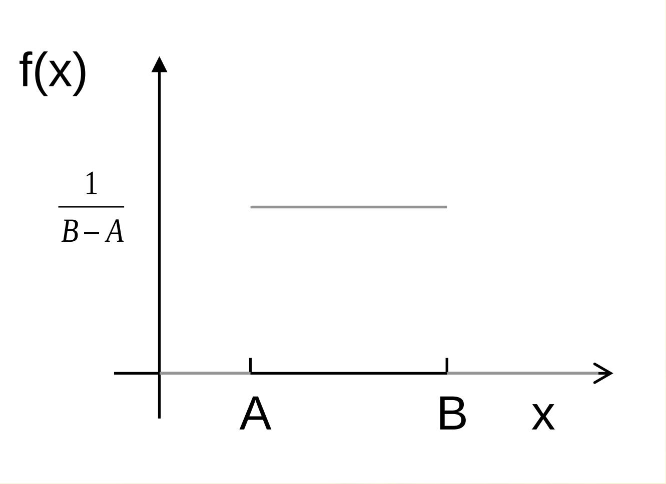
$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(y)dy$$

For each x, F(x) is the area under the density curve to the left of x as follows



- **Example 4.6**

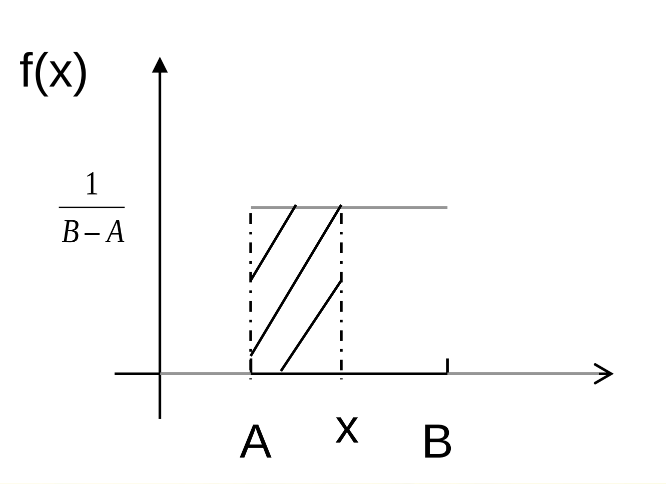
  Let X, the thickness of a certain metal sheet, have a uniform distribution on $[A, B]$. The density function is shown as follows.



What is the cdf?

Example 4.6 (Cont')

Solution:



For $x < A$, $F(x) = 0$, since there is no area under the graph of the density function to the left of such an x.

For $x \geq B$, $F(x) = 1$, since all the area is accumulated to the left of such an x.

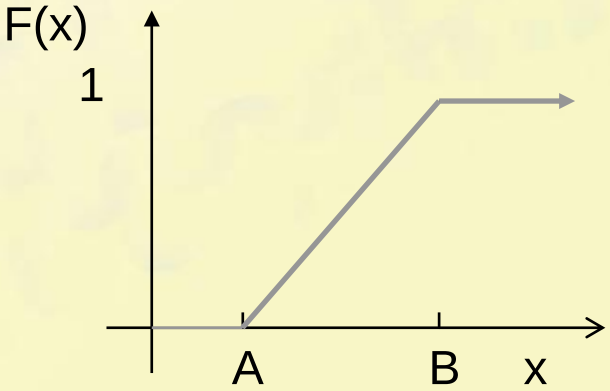
Example 4.6 (Cont')

For A ≤X≤ B

$$F(x) = \int_{-\infty}^{x} f(y)dy = \int_{A}^{x} \frac{1}{B-A} dy = \frac{1}{B-A} \cdot y \Big|_{y=A}^{y=x} = \frac{x-A}{B-A}$$

Therefore, the entire cdf is

$$F(x) = \begin{cases} 0 & x < A \\ \dfrac{x-A}{B-A} & A \le x < B \\ 1 & x \ge B \end{cases}$$



- **Using F(x) to compute probabilities**

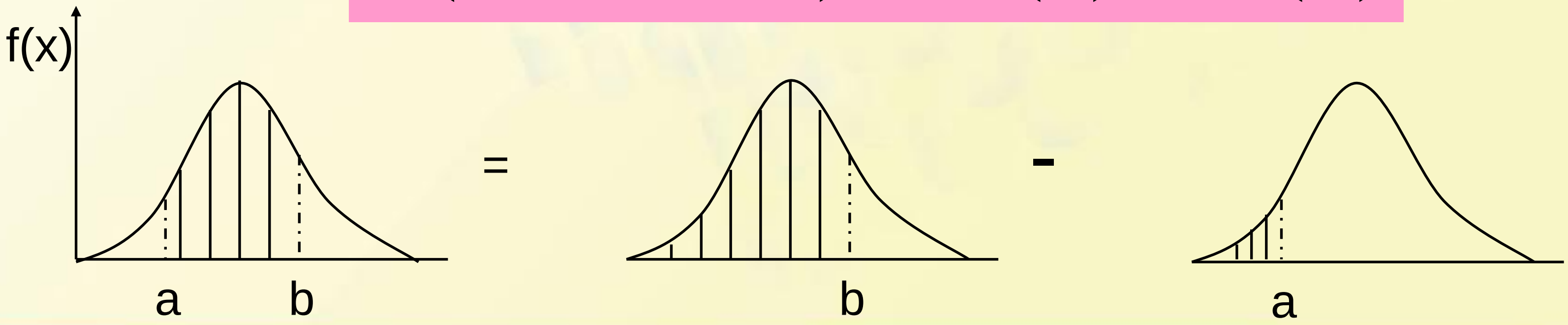
Let X be a continuous rv with pdf f(x) and cdf F(x).

Then for any number *a*

$$P(X > a) = 1 - F(a)$$

and for any two numbers *a* and *b* with *a<b*

$$P(a \le X \le b) = F(b) - F(a)$$



- **Example 4.7**

  Suppose the pdf of the magnitude X of a dynamic load on a bridge is given by

$$f(x) = \begin{cases} \frac{1}{8} + \frac{3}{8}x, & 0 \le x \le 2 \\ 0, & otherwise \end{cases}$$

find F(x), P(1≤X≤1.5) and P(X>1) ?

Solution:

$$F(x)=\int_{-\infty}^{x} f(y)dy=\int_{0}^{x}(\frac{1}{8}+\frac{3}{8}x)dy=\frac{x}{8}+\frac{3}{16}x^2$$

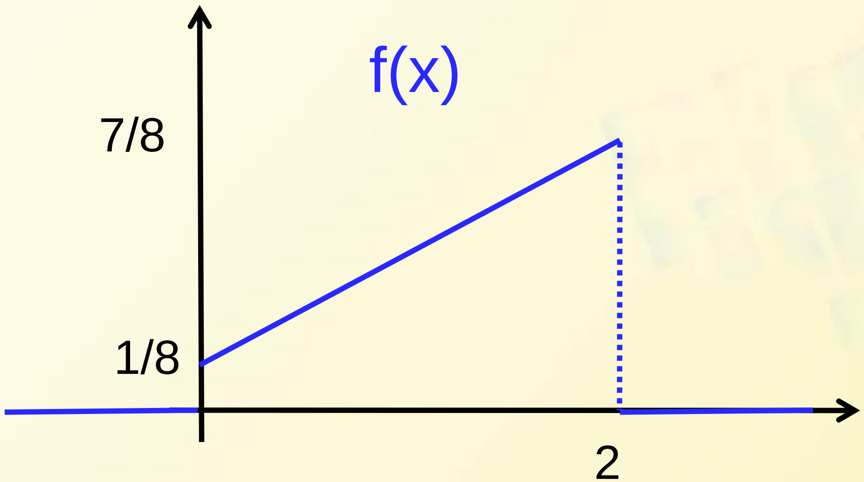
thus

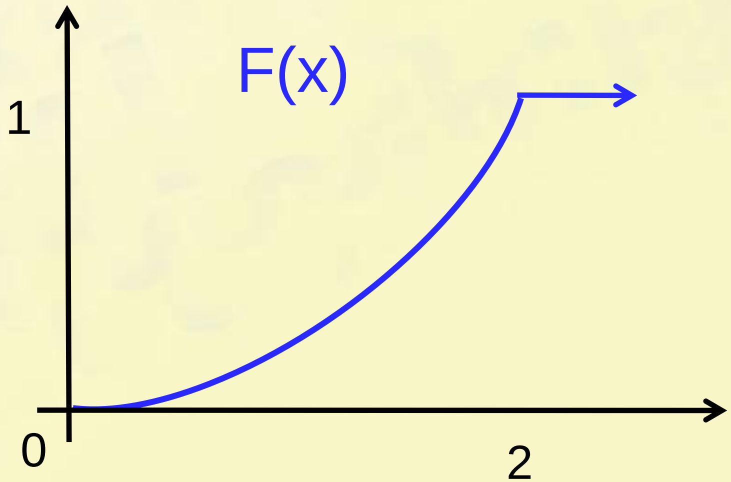
$$F(x)=\begin{cases} 0, x<0 \\ \int_{-\infty}^{x} f(y)dy=\int_{0}^{x}(\frac{1}{8}+\frac{3}{8}x)dy=\frac{x}{8}+\frac{3}{16}x^2, x\in[0,2] \\ 1, x>2 \end{cases}$$

- **Example 4.7 (Cont')**

$$P(1 \le X \le 1.5) = F(1.5) - F(1) = 0.297$$

$$P(X > 1) = 1 - F(X = 1) = 0.688$$





- **Obtaining f(x) from F(x)**

If X is a continuous rv with pdf f(x) and cdf F(x), then at every x at which the derivative F'(x) exists, F'(x)=f(x)

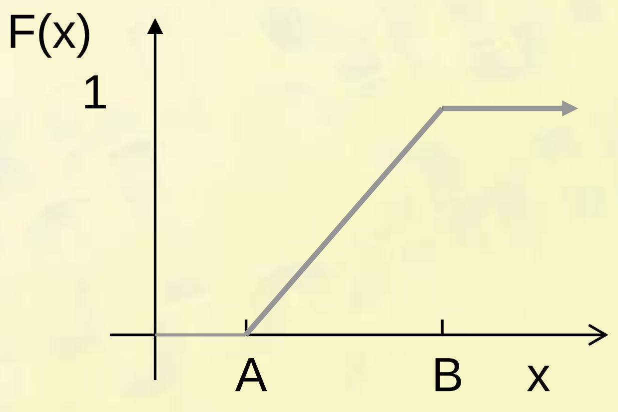
$$f(x) \Longrightarrow F(x) \qquad F(x) = P(X \le x) = \int_{-\infty}^{x} f(y)dy$$

$$F(x) \Longrightarrow f(x) \qquad f(x) = F'(x) = \left(\int_{-\infty}^{x} f(y)dy\right)'$$

Example 4.8 (Ex. 4.6 Cont')

When X has a uniform distribution, F(x) is differentiable except at x=A and x=B, where the graph of F(x) has sharp corners. (See the following Figure)

$$F(x)=\begin{cases} 0 & x<A \\ \dfrac{x-A}{B-A} & A\le x<B \\ 1 & x\ge B \end{cases}$$



Now, given the F(x), what is the f(x)?

Solution:

For $x<A$, $F(x)=0$;

For $x>B$, $F(x)=1$, $F'(x)=0=f(x)$ for such x.

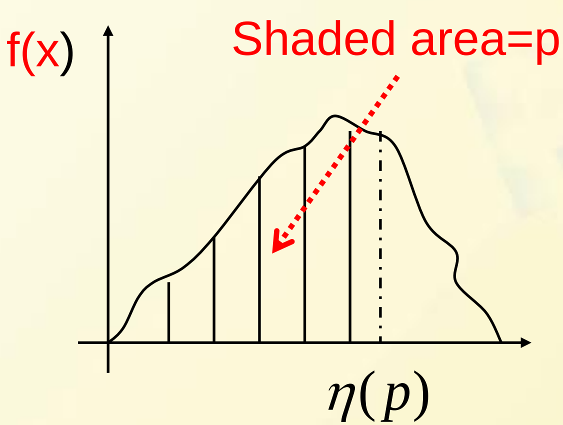
For $A<x<B$

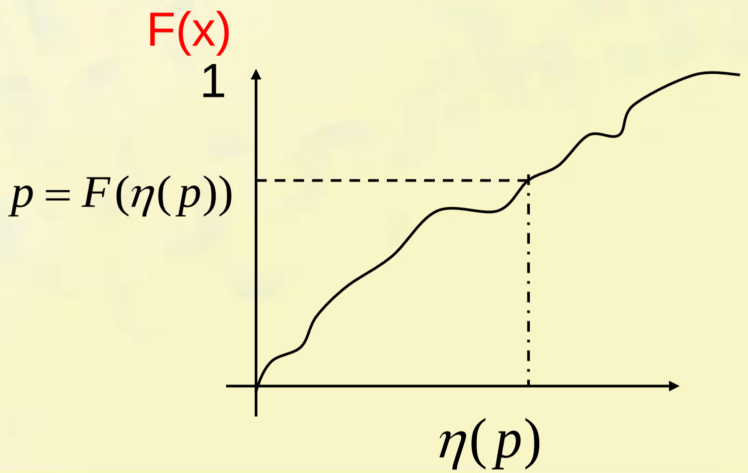
$$F'(x)=\frac{d}{dx}\left(\frac{x-A}{B-A}\right)=\frac{1}{B-A}=f(x)$$

▪ Percentiles of a Continuous Distribution

Let p be a number between 0 and 1. The (100p)th percentile of the distribution of a continuous rv X , denoted by η(p), is defined by

$$p = F(\eta(p)) = \int_{-\infty}^{\eta(p)} f(y)dy$$





- **Example 4.9**

The distribution of the amount of gravel (in tons) sold by a particular construction supply company in a given week is a continuous rv X with pdf

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

(A) What is the cdf?

(B) If p=0.5, η(p)=？

Solution:

(A) The cdf of sales for any x between 0 and 1 is

$$F(x) = \int_0^x \frac{3}{2}(1 - y^2)dy = \frac{3}{2}(y - \frac{y^3}{3})\Big|_{y=0}^{y=x} = \frac{3}{2}(x - \frac{x^3}{3})$$

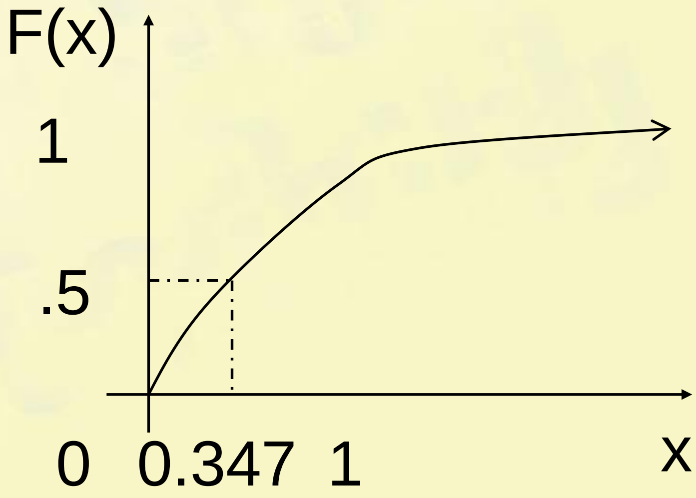
- **Example 4.9 (Cont')**

(B)

$$p = F(\eta(p)) = \frac{3}{2}\left[\eta(p) - \frac{(\eta(p))^3}{3}\right]$$

$$(\eta(p))^3 - 3\eta(p) + 2p = 0$$

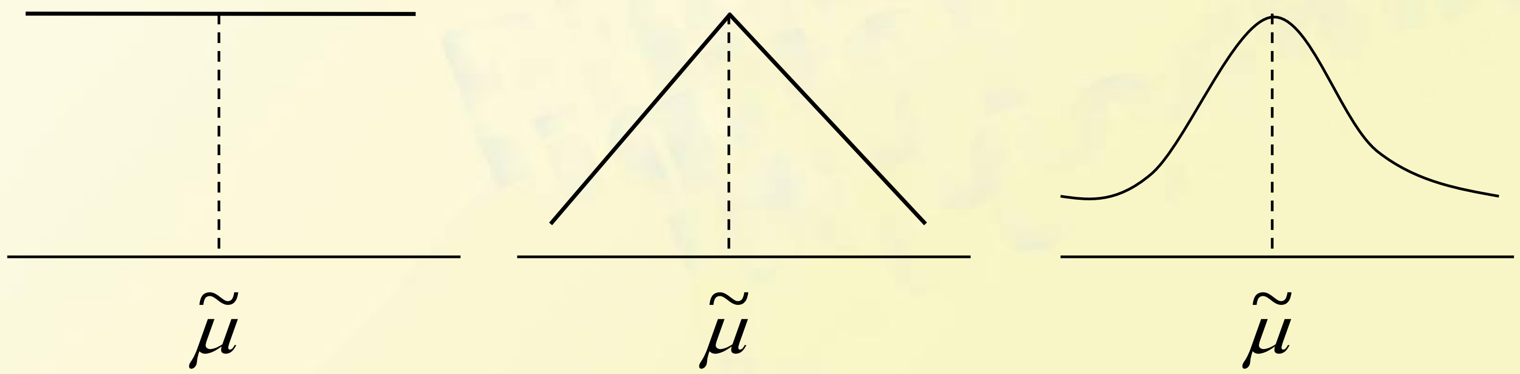
$$If \;\; p = 0.5, \eta(p) = 0.347$$



- **The median**

The median of a continuous distribution, denoted by $\tilde{\mu}$, is the 50^{th} percentile, so satisfies $0.5=F(\tilde{\mu})$, that is, half the area under the density curve is to the left of $\tilde{\mu}$ and half is to the right of $\tilde{\mu}$

Symmetric Distribution



- Expected/Mean Value

The expected/mean value of a continuous rv X with pdf f(x) is

$$\mu_X = E(X) = \int_{-\infty}^{\infty} xf(x)dx$$

$$\mu_X = E(X) = \sum_{x \in D} x \cdot p(x)$$

Discrete Case

- **Example 4.10 (Ex. 4.9 Cont')**

The pdf of weekly gravel sales X was

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

E(x)=?

Solution:

$$E(x) = \int_{-\infty}^{+\infty} x f(x) dx = \int_0^1 x \frac{3}{2}(1 - x^2) dx = \frac{3}{2}\left(\frac{x^2}{2} - \frac{x^4}{4}\right)\Big|_{x=0}^{x=1} = \frac{3}{8}$$

Expected value of a function

If X is a continuous rv with pdf $f(x)$ and $h(X)$ is any function of X, then

$$E\left[h(X)\right] = \mu_{h(X)} = \int_{-\infty}^{\infty} h(x)\, f\left(x\right)dx$$

$$\mu_{h(X)} = E(h(X)) = \sum_{x \in D} h(x) \cdot p(x)$$

Discrete Case

Example 4.11

Two species are competing in a region for control of a limited amount of a certain resource. Let X = the proportion of the resource controlled by species 1 and suppose X has pdf

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

which is a uniform distribution on [0,1]. Then the species that controls the majority of this resource controls the amount

$$h(X) = \max(X, 1-X) = \begin{cases} 1-X & \text{if } 0 \le x < \frac{1}{2} \\ X & \text{if } \frac{1}{2} \le X \le 1 \end{cases}$$

What is the E(h(X))?

Solution:

The expected amount controlled by the species having majority control is then

$$E[h(X)] = \int_{-\infty}^{\infty} \max(x, x-1) \cdot f(x) dx = \int_{0}^{1} \max(x, 1-x) \cdot 1 dx$$

$$= \int_{0}^{1/2} (1-x) \cdot 1 dx + \int_{1/2}^{1} x \cdot 1 dx = \frac{3}{4}$$

▪ The Variance

The variance of a continuous random variable X with pdf $f(x)$ and mean value μ is

$$\sigma_X{}^2 = V(X) = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx = E[(X-\mu)^2]$$

The standard deviation (SD) of X is

$$\sigma_X = \sqrt{V(X)}$$

- **Proposition**

$$E(aX+b)=aE(X)+b$$

$$V(X)=E(X^2)-[E(X)]^2$$

The Same Properties as Discrete Cases