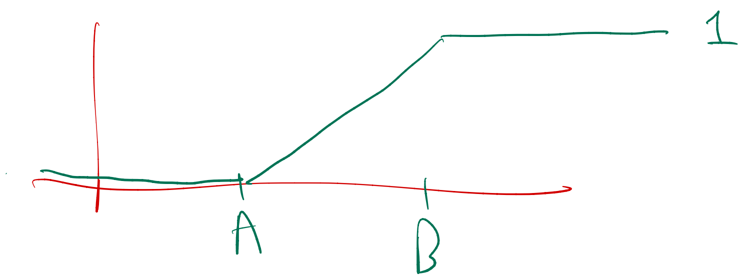
Lec 17

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. Finally, for $A \leq x \leq B$,

The entire cdf is

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

The graph of this cdf is



Using $F(x)$ to Compute Probabilities

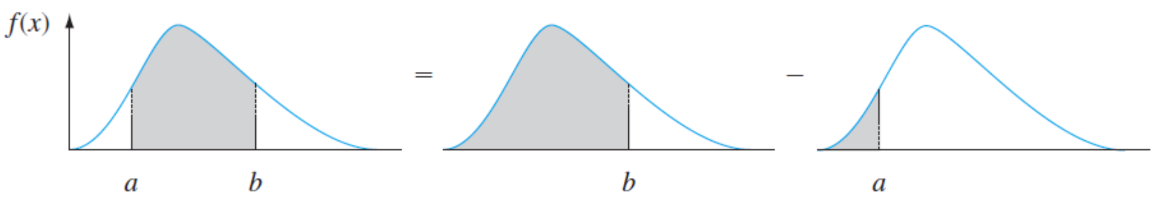
As for discrete random variables, probabilities of various intervals can be computed from a formula or table of $F(x)$.

> Let X be a continuous random variable with pdf $f(x)$ and cdf $F(x)$. Then for any number a,
>
> $$P(X > a) = 1 - P(X \leq a) = 1 - F(a)$$
>
> and for any two numbers a and b with $a < b$,
>
> $$P(a \leq X \leq b) = F(b) - F(a)$$

Figure below illustrates the second part of this proposition; the desired probability is the shaded area under the density curve between a and b, and it equals the difference between the two shaded cumulative areas.

$$P(X \leq b) - P(X \leq a)$$
$$= P(a < X \leq b)$$
$$= P(a \leq X \leq b)$$

NOTE: This is different from a discrete random variable (e.g., binomial or Poisson): $P(a \leq X \leq b) = F(b) - F(a-1)$ when a and b are integers.



Computing $P(a \leq X \leq b)$ from cumulative probabilities

Example 55. Suppose the pdf of the magnitude X of a dynamic load on a bridge (in newtons) is given by

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

$F(x) = 0 \quad x < 0$

$F(x) = 1 \quad x > 2$

For any number x between 0 and 2,

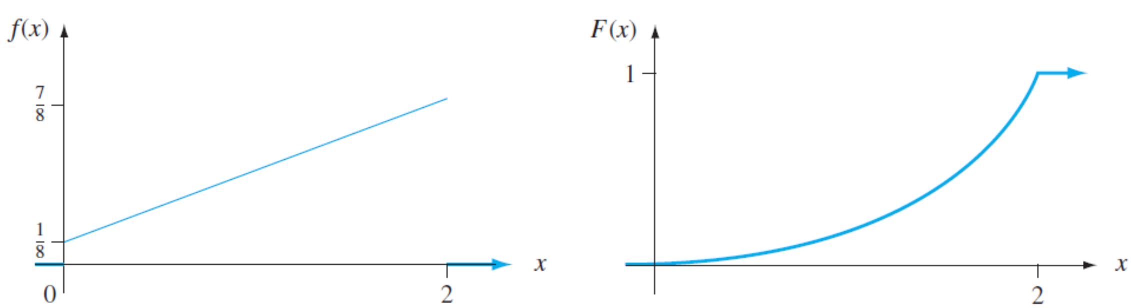
$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(s)\,ds$$

Thus in summary

$$= \int_0^x \left(\frac{1}{8} + \frac{3}{8}s\right) ds = \left(\frac{1}{8}s + \frac{3}{16}s^2\right)\Big|_0^x$$

$$= \frac{1}{8}x + \frac{3}{16}x^2$$

The graphs of $f(x)$ and $F(x)$ are shown in Figure below.



The pdf and cdf for Example 4.7

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{8}x + \frac{3}{16}x^2 & 0 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$

The probability that the load is between 1 and 1.5 is

$$P(1 \le X \le 1.5) = \int_1^{1.5} f(x)\,dx = \int_1^{1.5} \left(\frac{1}{8} + \frac{3}{8}x\right) dx = \ldots$$

$$\text{or} \quad = F(1.5) - F(1) = \ldots$$

The probability that the load exceeds 1 is

$$P(X \ge 1) = \int_1^{\infty} f(x)\,dx = \int_1^{2} \left(\frac{1}{8} + \frac{3}{8}x\right) dx$$

$$\text{or} \quad = 1 - F(1)$$

Obtaining $f(x)$ from $F(x)$

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

Recall: For X discrete, the pmf is obtained from the cdf by taking the difference between two $F(x)$ values.

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

This result is a consequence of the Fundamental Theorem of Calculus.

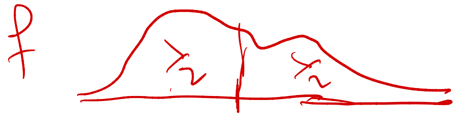
Example 56. (Example 54 continued) When X has a uniform distribution, $F(x)$ is differentiable except at $x = A$ and $x = B$, where the graph of $F(x)$ has sharp corners. Since $F(x) = 0$ for $x < A$ and $F(x) = 1$ for $x > B$, $F'(x) = 0 = f(x)$ for such x. For $A < x < B$,

unif(0,1)

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

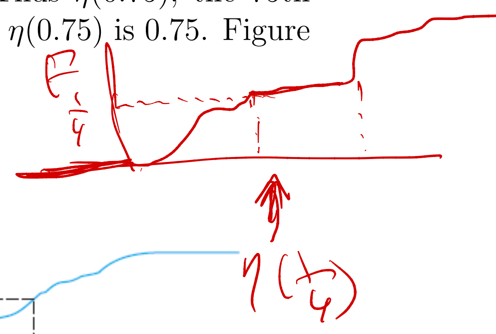
Percentiles of a Continuous Distribution

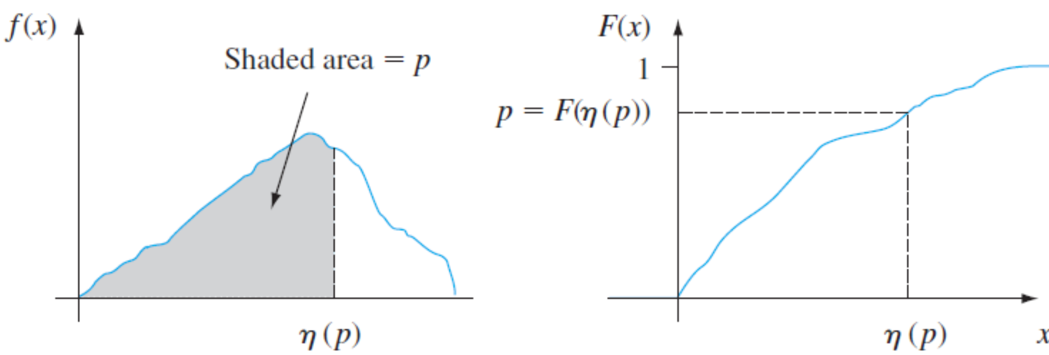
When we say that an individual's test score was at the 85th percentile of the population, we mean that 85% of all population scores were below that score and 15% were above.



> **Definition 18.** Let p be a number between 0 and 1. The **(100p)th percentile** of the distribution of a continuous random variable X, denoted by $\eta(p)$, is defined by
>
> $x\%$ to $\quad F(x) = p$

According to Definition 18, $\eta(p)$ is that value such that $100p\%$ of the area under the graph of $f(x)$ lies to the left of $\eta(p)$ and $100(1-p)\%$ lies to the right. Thus $\eta(0.75)$, the 75th percentile, is such that the area under the graph of $f(x)$ to the left of $\eta(0.75)$ is 0.75. Figure below illustrates the definition.





The (100p)th percentile of a continuous distribution

Example 57. The distribution of the amount of gravel (in tons) sold by a particular construction supply company in a given week is a continuous random variable X with pdf

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

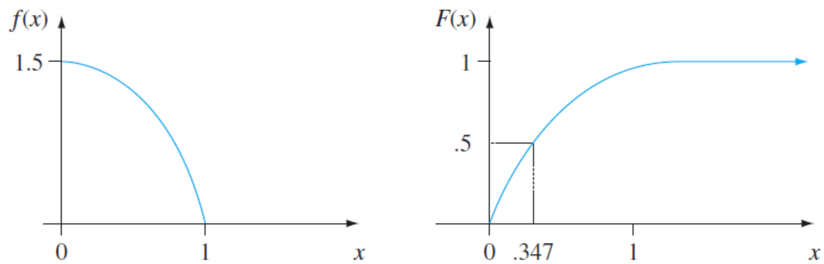
$$\int_0^x \frac{3}{2}(1-s^2)\,ds$$

$$\left(\frac{3}{2}s - \frac{1}{2}s^3\right)\Big|_0^x$$

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

$$F(x) = \begin{cases} 0 & x < 0 \\ \frac{3}{2}x - \frac{1}{2}x^3 & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

The graphs of both $f(x)$ and $F(x)$ appear in Figure below.



The $(100p)$th percentile of this distribution satisfies the equation

solve: $\frac{3}{2}x - \frac{1}{2}x^3 = p$

For the 50th percentile, $p = 0.5$, and the equation to be solved is

$$\frac{3}{2}x - \frac{1}{2}x^3 = \frac{1}{2}$$

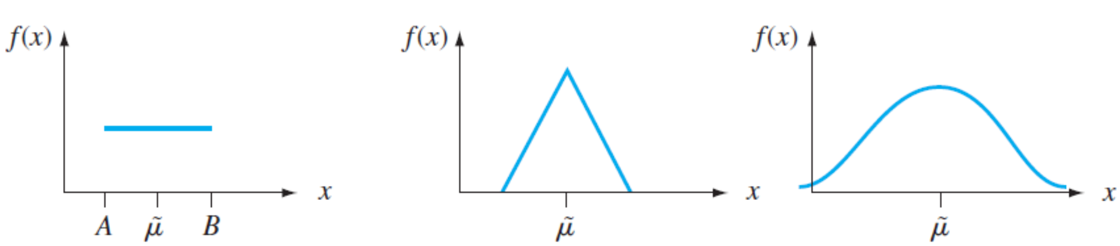
the solution is

$p = 0.347$

Which means if the distribution remains the same from week to week, then in the long run 50% of all weeks will result in sales of less than 0.347 ton and 50% in more than 0.347 ton.

Definition 19. The median of a continuous distribution, denoted by $\widetilde{\mu}$, is the 50th percentile, so $\widetilde{\mu}$, satisfies $0.5 = F(\widetilde{\mu})$. That is, half the area under the density curve is to the left of $\widetilde{\mu}$ and half is to the right of $\widetilde{\mu}$.

A continuous distribution whose pdf is **symmetric** - the graph of the pdf to the left of some point is a mirror image of the graph to the right of that point - has median $\widetilde{\mu}$, equal to the point of symmetry, since half the area under the curve lies to either side of this point. Figure below gives several examples.



Medians of symmetric distributions