Recall: $Var(aX + b) = a^2 Var(X) + 0$

↳ b: constants don't vary

$$sd(aX + b) = |a|\sqrt{Var(X)}$$

<u>Continuous random variables</u> take on values on the real number line

- No longer countable
- Infinite number of possibilities X=x
- P(X=a) = 0 for each a, since the probability of getting a specific value among an infinite number of possibilities is extremely small

Instead, we specify probability over *intervals -> area under the curve*

- Discrete probabilities involve summing over all values X=x
- Continuous probabilities involve taking the area under the curve, which is integration, which is summing over infinitesimally small values

We use the <u>cumulative distribution function</u> to describe the distribution of a continuous RV

- F(x) = P(X ≤ x)

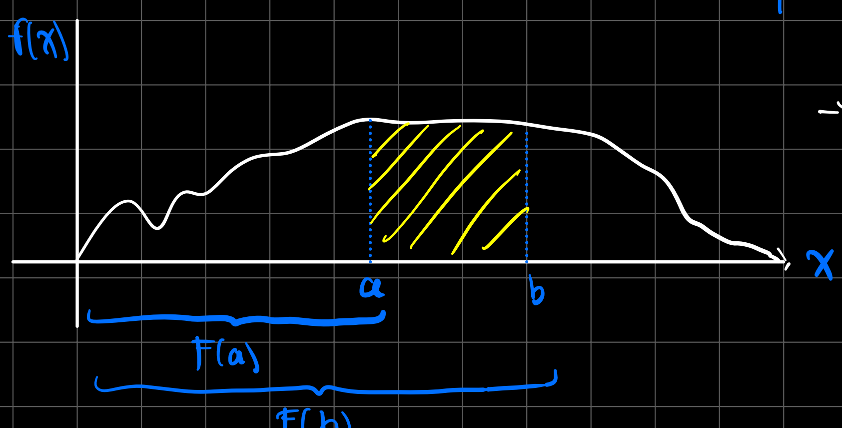
Satisfies the following properties:

- Defined for all real numbers x

$$\lim_{x \to -\infty} F(x) = 0 \qquad \lim_{x \to \infty} F(x) = 1$$

- F(x) is non-decreasing
- P(a < x ≤ b) = F(b) - F(a)

can also be $a < x < b$ or $a \leq x \leq b$ ... equality doesn't matter



$$\rightarrow P(a < X < b) = \int_a^b f(x)\,dx$$

$$= F(b) - F(a)$$

Probability Density Function

Suppose we take a short interval of values Δx:

$P(x \le X \le x + \Delta x) = F(x + \Delta x) - F(x)$

Relationship between the PDF and the CDF:

$$\underbrace{\lim_{\Delta x \to 0} \frac{P(x \le X \le x + \Delta x)}{\Delta x}}_{\text{pdf of } x} = \underbrace{\lim_{\Delta x \to 0} \frac{F(x + \Delta x) - F(x)}{\Delta x}}_{\text{def'n of derivative}}$$

$$\Rightarrow f(x) = F'(x)$$

The PDF is f(x), and:

$$F(x) = \int_{-\infty}^{x} f(u)\, du \qquad \int_{-\infty}^{\infty} f(x)\, dx = 1$$

f(x): area under x

F(X): summed area under x for all x≤X — CDF

$$P(X = a) = \int_a^a f(x)\, dx = F(a) - F(a) = 0$$

$$\text{So } f(x) \neq P(X = x)$$

Note that the p.d.f., $f(x)$ represents the likelihood of (small intervals around) different x-values. To see this, note that for small Δx

$$P\left(x - \frac{\Delta x}{2} \le X \le x + \frac{\Delta x}{2}\right)$$
$$= F\left(x + \frac{\Delta x}{2}\right) - F\left(x - \frac{\Delta x}{2}\right) \approx f(x)\Delta x$$

Let X be a r.v. with c.d.f.

$$F(x) = \begin{cases} 0 & \text{for } x \leq 0 \\ \frac{x}{4} & \text{for } 0 < x \leq 4 \\ 1 & \text{for } x > 4 \end{cases}$$

a) Find the p.d.f. $f(x)$.
b) Solve for the 90th percentile (i.e. solve for the value, x, such that the area under the curve to the left of that value is 0.9. Maybe call this value x_{90}).

(a) Since $f(x) = F'(x)$:

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

(b) We want $F(x) = 0.9$

Since $F(x) = \frac{x}{4}$ for $0 < x \leq 4$ (0.9 is in this interval):

$$\frac{x_{90}}{4} = 0.9 \rightarrow x_{90} = 3.6$$

Or $P(X \leq 3.6) = 0.9$

Example: Suppose X is a continuous r.v. having p.d.f.

$$f(x) = \frac{1}{4} \text{ for } 0 < x < 4$$

and c.d.f.

$$F(x) = \frac{x}{4} \text{ for } 0 \leq x \leq 4.$$

Let $Y = 1/X$.
Find the p.d.f. of Y.

(1) Write the CDF of Y as a function of X:

$$F_y(Y) = P(Y \leq y) = P(1/x \leq y) = P(X \geq \tfrac{1}{y}) = 1 - P(X < \tfrac{1}{y})$$

(2) Use F_x to find F_y:

$$\Rightarrow F_y(Y) = 1 - F_x\left(\frac{1}{y}\right)$$

Since $F(x) = \frac{x}{4}$ for $0 \le x \le 4$:

$$F_y(Y) = 1 - \frac{1}{4y}$$

Differentiate to get the PDF of y

Expectation, Mean, and Variance of Continuous Distributions

Recall:

$$E(g(x)) = \sum_{\forall x} g(x) \cdot f(x) \qquad 0 \le f(x) \le 1$$

$$x \cdot f(x)$$

So for continuous distributions:

$$E(g(x)) = \int_{-\infty}^{\infty} g(x) \cdot f(x)\, dx$$

And the mean is:

$$\mu = E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

Example: If X is a continuous r.v. having p.d.f.

$$f(x) = \frac{1}{4} \text{ for } 0 < x < 4,$$

Calculate $E(X)$ and $Var(X)$.

$$E(x) = \int_0^4 x \cdot \frac{1}{4}\, dx = 2$$

$$Var(X) = E(X^2) - \mu^2$$

$$= E(X^2) - [E(X)]^2 \quad : \text{ just integrate}$$

Discrete distribution -> sum

Continuous distribution -> integrate