Lecture 10 - PDFs

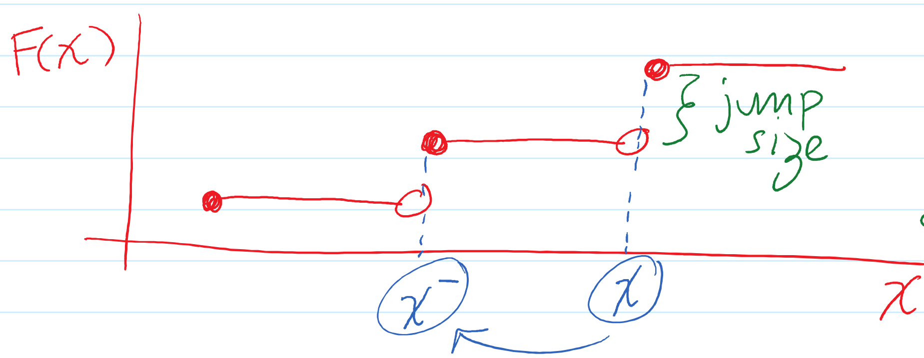
Thursday, October 7, 2021 1:57 PM

PMF: $f(x) = \mathbb{P}(X = x)$

CDF: $F(x) = \mathbb{P}(X \le x)$

$$\boxed{F(x) = \sum_{i \le x} f(i)}$$ ← from f to F

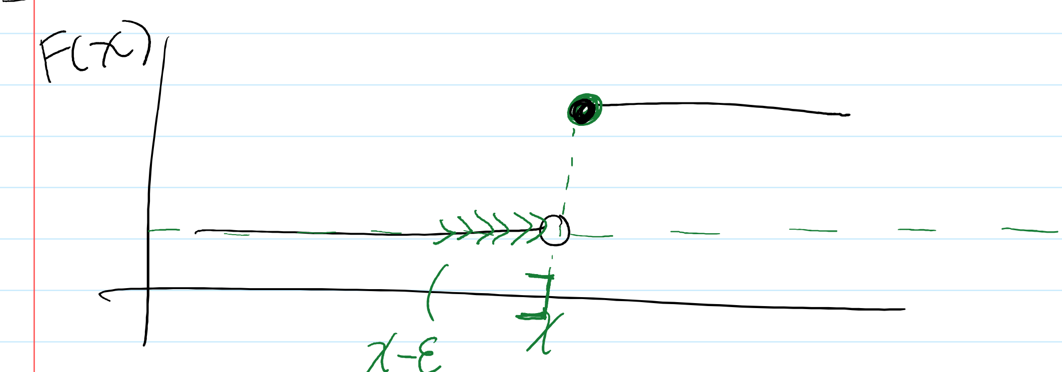
$$\mathbb{P}(a < X \le b) = F(b) - F(a)$$



next lowest point below x

$\mathbb{P}(X = x)$
$= \mathbb{P}(x^- < X \le x)$
$= F(x) - F(x^-)$
$=$ jump at x

from F to f

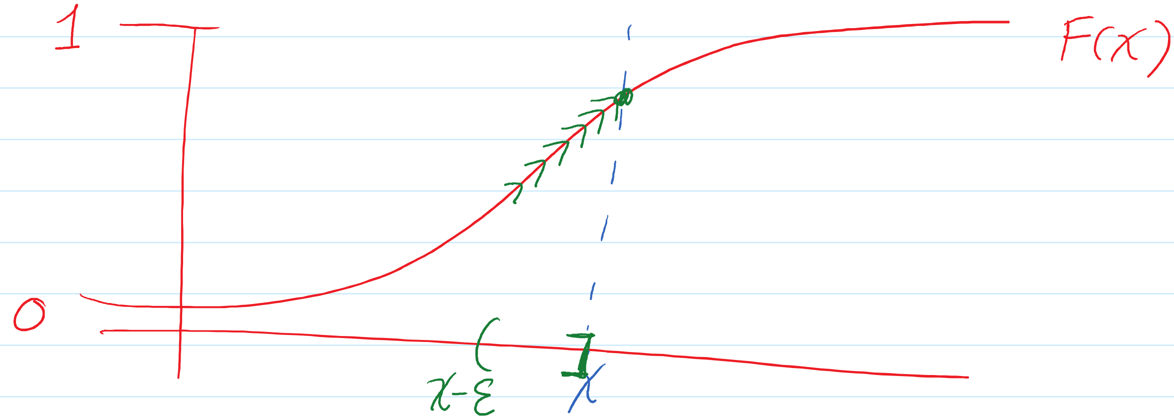


$$\lim_{\varepsilon \downarrow 0} \mathbb{P}(x - \varepsilon < X \le x) = \mathbb{P}(X = x) = f(x)$$

$$= \lim_{\varepsilon \downarrow 0} F(x) - F(x - \varepsilon)$$

$=$ jump size

What about continuous RVs?



$$\lim_{\varepsilon \downarrow 0} P(x-\varepsilon < X \le x) = P(X=x)$$

$$= \lim_{\varepsilon \downarrow 0} F(x) - \underbrace{F(x-\varepsilon)}_{\text{b/c } F \text{ is continuous}} \qquad F(x-\varepsilon) \to F(x)$$

$$= F(x) - F(x)$$

$$= 0$$

Punchline: this PMF formulation isn't useful for cts RVs

Can we create a continuous analog to the PMF?

Want: $F(x) = \sum_{i \le x} f(i)$ [for discrete]

Defn: Probability Density Function (PDF)

Analog of PMF for cts RVs.

The PDF for a cts RV is a function
$f: \mathbb{R} \to \mathbb{R}$ defined $\forall x \in \mathbb{R}$ as the function for which

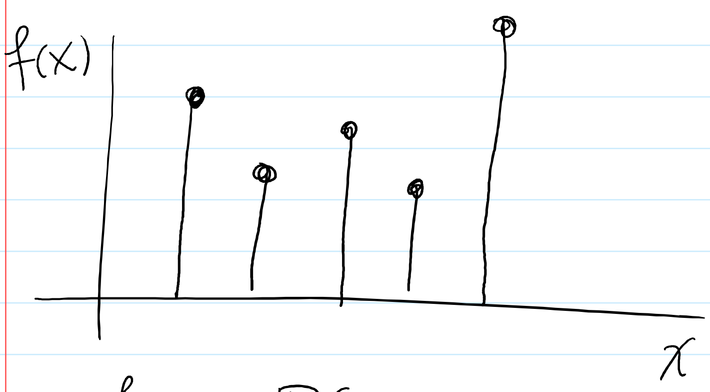
$$F(x) = \int_{-\infty}^{x} f(t)\,dt.$$

note by the Fundamental Theorem of Calculus

$$\frac{dF}{dx} = \frac{d}{dx}\int_{-\infty}^{x} f(t)\,dt = f(x)$$

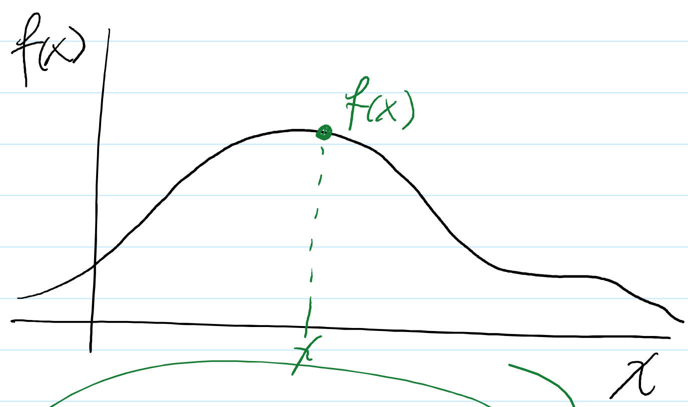
PDF = derivative of the CDF.

discrete PMF



$f(x) = \mathbb{P}(X = x)$

continuous PDF



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

Properties of PDF

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

$$= \int_{-\infty}^{b} f(t)\,dt - \int_{-\infty}^{a} f(t)\,dt$$

$$\boxed{= \int_a^b f(t)\,dt}$$

we said $P(X=a) = P(X=b) = 0$

& $P(a < X \le b)$
$$\left.\begin{aligned} &= P(a \le X \le b) \\ &= P(a \le X < b) \\ &= P(a < X < b) \end{aligned}\right\}$$ note: only for cts RVs

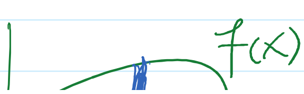
<u>More general fact</u>:

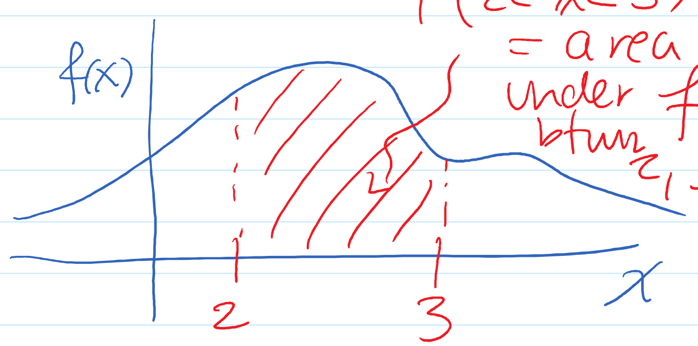
(discrete): $P(X \in A) = \sum_{x \in A} f(x)$

(continuous): $P(X \in A) = \int_A f(x)\,dx$

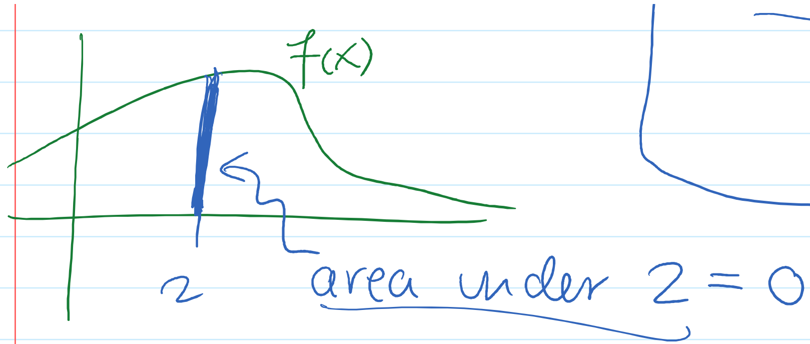
<u>Ex.</u> $P(X \in (2,3]) = \int_2^3 f(x)\,dx$

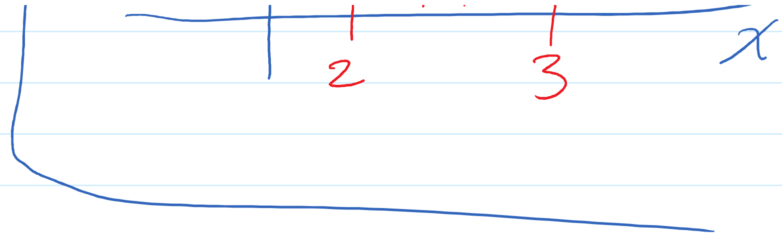
<u>Ex.</u> $P(X \in \{2\})$
$= P(X=2) = 0$





$P(2 < X < 3)$
$=$ area under f btwn $2, 3$





Ex. $F(x) = \dfrac{1}{1+e^{-x}}$

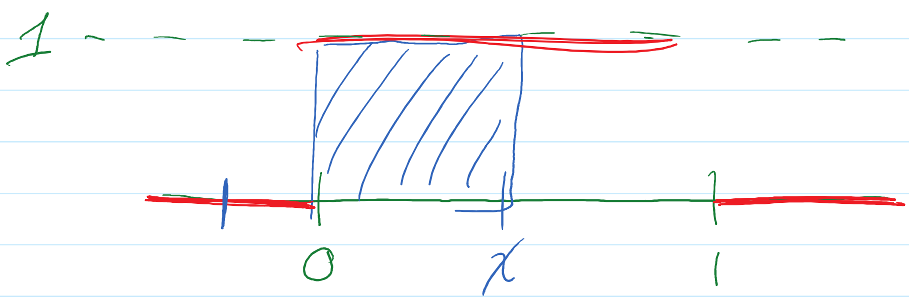
what is the associated PDF?

$$f(x) = \frac{d}{dx}F(x) = \ldots = \frac{e^{-x}}{(1+e^{-x})^2}.$$

Ex. Continuous Uniform Distribution (on $(0,1)$)

$X \sim U(0,1)$.

means $f(x) = \begin{cases} 1, & 0 < x < 1 \\ 0 & \text{else} \end{cases}$



What is the CDF? $F(x) = \int_{-\infty}^{x} f(t)\,dt$

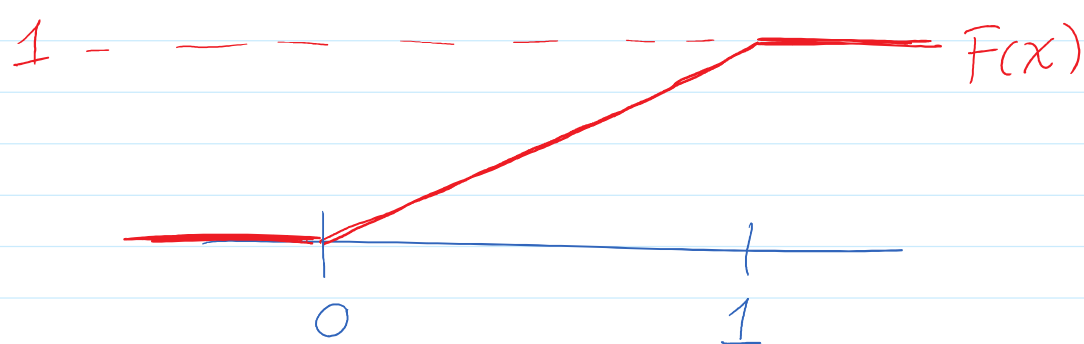
What is the CDF? $F(x)=\int_{-\infty}^{x} f(t)\,dt$

If $x<0$ then $F(x)=\int_{-\infty}^{x} 0\,dt = 0$

If $0<x<1$ then $F(x)=\int_{-\infty}^{x} f(t)\,dt = \int_{0}^{x} 1\,dt = x$

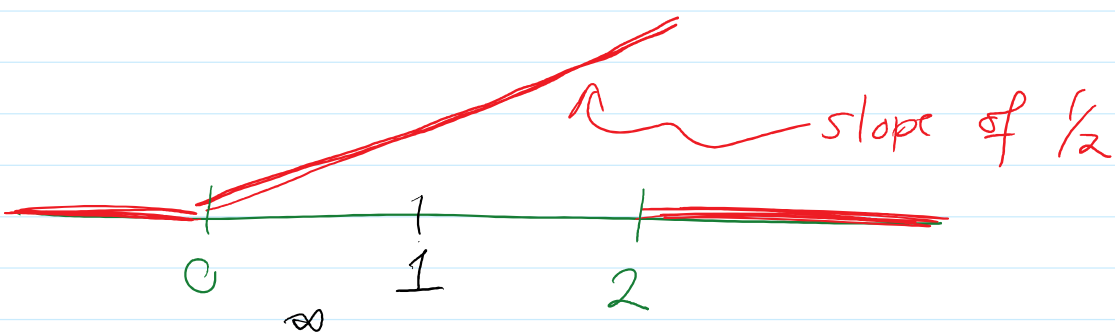
if $x>1$ then $F(x)=\int_{-\infty}^{\infty} f(t)\,dt = \int_{0}^{1} 1\,dt = 1$

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



Ex. Let

$$f(x)=\begin{cases} x/2 & 0<x<2 \\ 0 & \text{else} \end{cases}$$

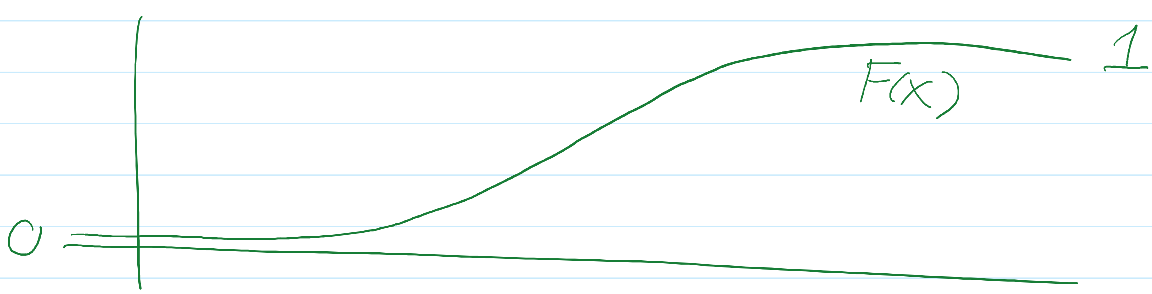


$$P(X>1) = \int_1^\infty f(t)\,dt = \int_1^2 t/2\,dt$$

$$= \frac{t^2}{4}\Big|_1^2 = \frac{4-1}{4} = 3/4$$

Ex. Let

$$F(x) = \begin{cases} 1-e^{-x} & \text{for } x>0 \\ 0 & \text{else} \end{cases}$$



Q: $P(1<X<2)$?

Way 1: $P(1<X<2) = F(2) - F(1)$

$$= (1-e^{-2}) - (1-e^{-1})$$

$$= e^{-1} - e^{-2}$$

<u>way 2:</u> $f(x) = \frac{dF}{dx} = \frac{d}{dx}\left(1 - e^{-x}\right) = e^{-x}$

and

$$P(1 < X < 2) = \int_1^2 e^{-x}\,dx = -e^{-x}\Big|_1^2$$

$$= -e^{-2} - (-e^{-1})$$

$$= e^{-1} - e^{-2}$$

<u>Theorem: PMF/PDF characterization</u>

A function f is the PMF/PDF of some RV iff

① $f(x) \geq 0 \quad \forall x \in \mathbb{R}$

② (discrete) $\sum_{x \in \mathbb{R}} f(x) = 1$

(continuous) $\int_{\mathbb{R}} f(x)\,dx = 1$

<u>aside:</u> <u>cts case:</u>

$$P(X \in A) = \int_A \underbrace{f(t)}_{\text{should be } \geq 0 \text{ so integral is}}\,dt \geq 0$$

$$1 = P(S) = P(X \in \mathbb{R}) = \int_{\mathbb{R}} f(t)\,dt$$

Fact: If $g(x) \geq 0$ and $\int_{\mathbb{R}} g(x)\,dx = c < \infty$

If $f(x) = \frac{g(x)}{c}$ then f is a PDF

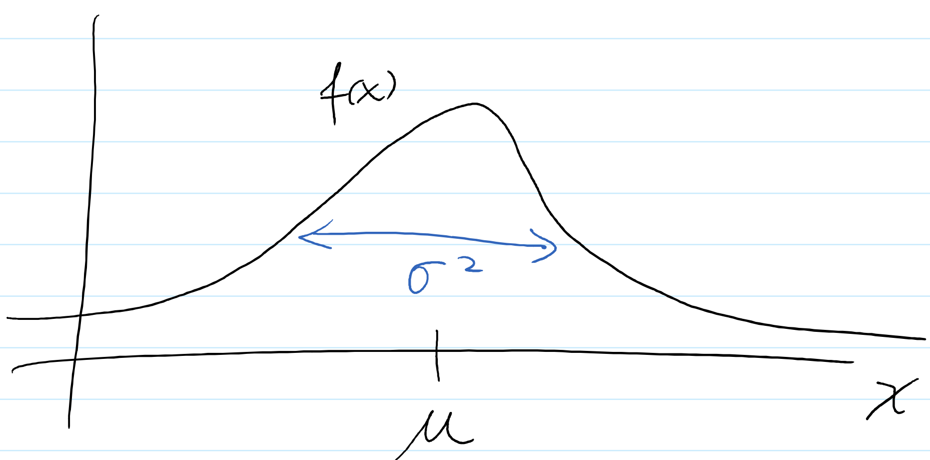
1. $f(x) \geq 0$
2. $\int f(x)\,dx = \int \frac{1}{c} g(x)\,dx = \frac{1}{c} c = 1$.

Ex. Normal Distribution (Gaussian)

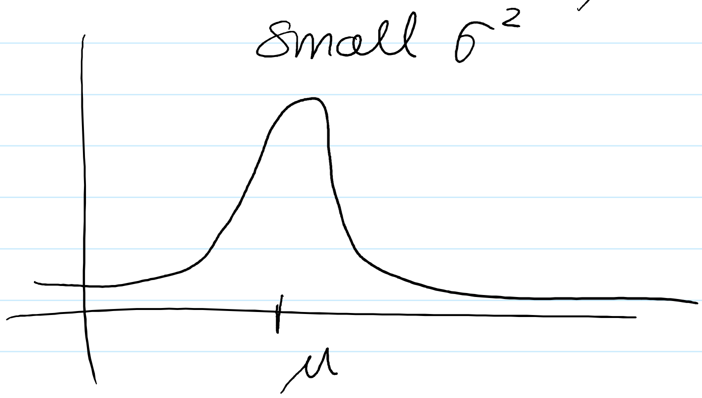
notation:

$$X \sim N(\mu, \sigma^2)$$

mean: $\mu \in \mathbb{R}$ variance: $\sigma^2 > 0$



Small σ^2 Large σ^2



Special case: standard normal

$$\mu = 0 \quad \text{and} \quad \sigma^2 = 1$$

$$X \sim N(0,1).$$

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) \quad \forall x \in \mathbb{R}$$

$\exp(a) = e^a$

Q: Is this a valid PDF?

(1) $f(x) \geq 0$ ✓

(2) $\int_{\mathbb{R}} f(x)\,dx = 1$

$$\underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx}_{I} \stackrel{?}{=} 1$$

Want: $I = 1. \Leftrightarrow I^2 = 1.$

∞ ∞

Want: $I^2 = 1 \Leftrightarrow I = 1$.

$$I^2 = I \cdot I = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) dx \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}y^2\right) dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{1}{2\pi} \exp\left(-\frac{1}{2}x^2\right) \exp\left(-\frac{1}{2}y^2\right) dx\, dy$$

$e^a e^b = e^{a+b}$

$$= \iint_{\mathbb{R}^2} \frac{1}{2\pi} \exp\left(-\frac{1}{2}(x^2+y^2)\right) dx\, dy$$

Polar coordinates.