5) Continuous Random Variables

Covers cases where the image is uncountable.

In discrete random variables, the distribution functions $F_X(x)$ where discontinuous. They were step functions.

In continuous random variables; the distribution function is differentiable $\Rightarrow$ continuous

Continuous random variables are charecterized by a property of their distribution function

Defn 5.1: We call a random variable X continuous if its distribution function F_X can be written as

$$\boxed{F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f(s)\,ds \qquad \forall x \in \mathbb{R}}$$

for some function $f_X : \mathbb{R} \to \mathbb{R}$. In this case we say f_X is the density function of X.

The fundemental theorem of calculus under some mild regularity conditions, that for each $x \in \mathbb{R}$

$$\boxed{\frac{d}{dx} F_X(x) = \frac{d}{dx} \int_{-\infty}^{x} f_X(s)\,ds = f_X(x)}$$

Thus:

- density function to distribution function requires integration.
- distribution function to to density function requires differentiation.

For calculating probabilities of events, involving random variables, density functions have for continuous random variables the same function mass functions have for discrete random variables.

Density functions have similar properties to m1 and m2 of Thm 4.5.

Theorem 5.2: Let X be a continuous random variable. Then its density function f_X satisfies

(d1) $$f_X(x) \geq 0 \quad \forall x \in \mathbb{R}$$

(d2) $$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$$

Conversely any real function f_X satisfying (d1) and (d2) is the density function of some random variable.

Property (d1), the non-negativity of $f_X(x)$ ensures $F_X(x)$ is an increasing function required by Theorem 4.8 ii

Also from Thm 4.8 iii we obtain

$$1 = \lim_{x \to \infty} F_X(x) = \lim_{x \to \infty} \int_{-\infty}^{x} f_X(x)$$

$$= \int_{-\infty}^{\infty} f_X(x)$$

which gives (d2).

Theorem 5.3: If X is a continuous random variable with density function f_X then

$\forall a, b \in \mathbb{R}$ with $a \leq b$

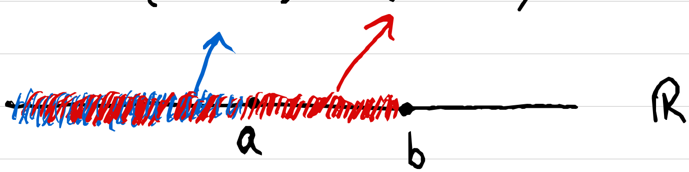
$$\boxed{P(A < x \leq b) = \int_a^b f_X(x)\,dx}$$

proof: We note that

$$\{a \leq X \leq b\} = \{X \leq b\} \cap \{X > a\}$$
$$= \{X \leq b\} \cap \{X \leq a\}^c$$

Furthermore

$$\{X \leq a\} \subseteq \{X \leq b\} \qquad (*1)$$



Hence we can now use the fact:

If A and B are events, with $A \subseteq B$ then

$$P(B \cap A^c) = P(B) - P(A)$$

$$P(a < X \leq b) = P(\{X \leq b\} \cap \{X \leq a\}^c)$$

$$= P(X \leq b) - P(X \leq a) \quad (\text{by } *1)$$

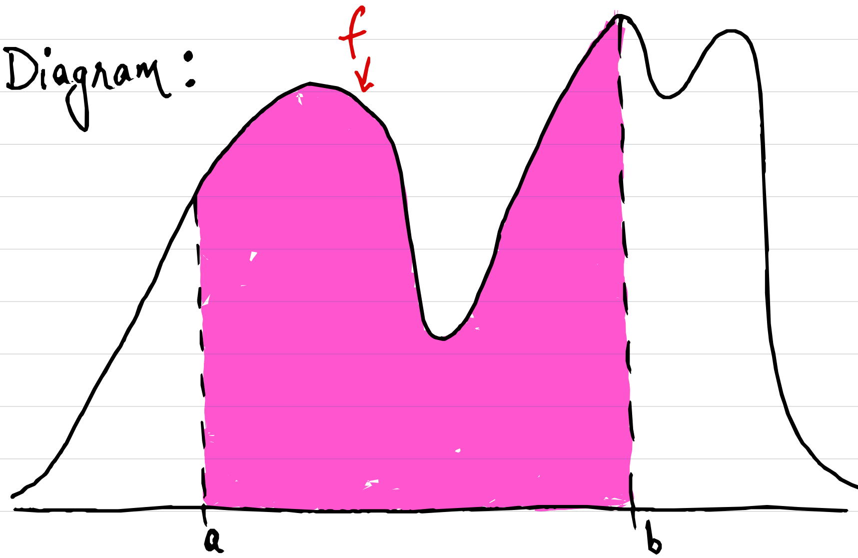
$$= F_X(b) - F_X(a)$$

$$= \int_{-\infty}^{b} f_X(x)\,dx - \int_{-\infty}^{a} f_X(x)\,dx$$

by Def 5.1

$$= \int_{a}^{b} f_X(x)\,dx$$

Diagram:



So the theorem says that $P(a < X \leq b)$ can be calculated by calculating area under density function (integration) between points a and b.

If the interval gets smaller and smaller, the probability will go to 0. i.e.
For any $\varepsilon > 0$,

$$P(a-\varepsilon \leq X \leq a+\varepsilon) = \int_{a-\varepsilon}^{a+\varepsilon} f_X(x)\,dx$$

and as $\varepsilon \to 0$, $P(a-\varepsilon \leq X \leq a+\varepsilon) \to P(a \leq X \leq a)$
$= P(X=a)$

and as $\varepsilon \to 0$, $P(X=a) = 0$.

This implies that for continuous random variables, you can be careless about precise form of inequalities

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

From Thm 5.3 we can see that

$$\boxed{P(a \le X \le b) = P(X \le b) - P(X \le b)}$$

or

$$\boxed{\int_a^b f_X(x) = \int_{-\infty}^{b} f_X(x) - \int_{-\infty}^{a} f_X(x)}$$

Theorem 5.4 If X is a continuous random variable, then $\forall x \in \mathbb{R}$

$$P(X=x) = 0$$

As a consequence $P(X=a) = 0 = P(X=b)$

So in Thm 5.3, it does not matter whether we use weak inequalities ($\leq$) or strict inequalities ($<$)

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

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

$$= \int_a^b f_X(x)\,dx$$

5.2 Frequently used Continuous probability distribution.

Defn 5.5: Uniform Distribution:

We say that a continuous random variable X has the uniform distribution on $[a,b]$ and write

$$\boxed{X \sim U(a,b)}$$

if the density function is

$$\boxed{f_X(x) = \begin{cases} \frac{1}{b-a} & x \in [a,b] \\ 0 & x \notin [a,b] \end{cases}}$$

The above is indeed a density function since $f(x) \geq 0 \ \forall x$ and

$$\int_{-\infty}^{\infty} f_X(x) = \frac{1}{b-a}\int_a^b dx = \left(\frac{1}{b-a}\right)[x]_a^b$$

$$= 1$$

For $x \in [a,b]$ the distribution function is given by

$$F_X(x) = \int_{-\infty}^{x} f_X(s)\, ds$$

$$= \int_{-\infty}^{a} f_X(s) ds + \int_{a}^{x} \frac{1}{b-a} ds$$

$$= 0 + \left[\frac{s}{b-a} \right]_a^x$$

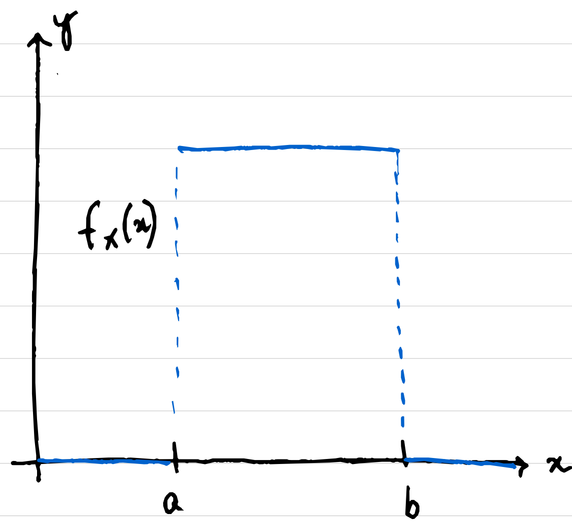
$$= \frac{x-a}{b-a}$$

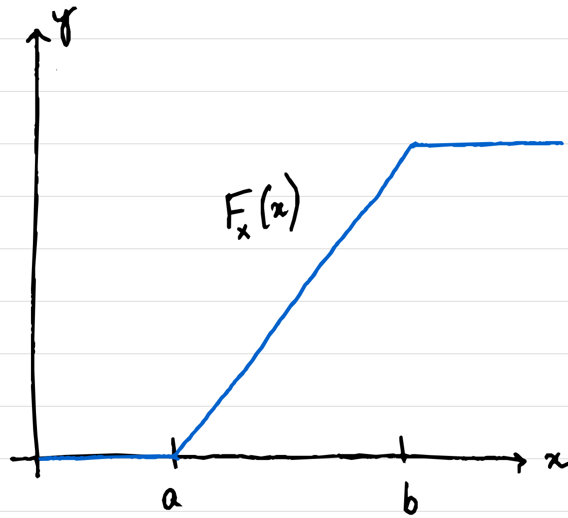
Thus full specification of <u>distribution fn</u>:

$$F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a,b] \\ 1 & x > b \end{cases}$$

Uniform distribution is used for example when we talk about choosing a number at random from an interval $[a,b]$.
An informal but correct description would be to say that a random number is equally likely to take any value in the interval.

Graphs:





Def 1.5.6 **Exponential Distribution**

We say that the continuous random variable X has the <u>exponential distribution</u> with parameter $\lambda > 0$ and write

$$\boxed{X \sim Exp(\lambda)}$$

if the density function of X is

$$\boxed{f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}}$$

This is a density function since $\underline{f_X(x) \geq 0}$ $\underline{\text{for all } x \in \mathbb{R}}$
(notice that this $\underline{\text{requires } \lambda \geq 0}$)
and

$$\int_{-\infty}^{\infty} f_X(x)\, dx = \int_{-\infty}^{0} f_X(x)\,dx + \int_{0}^{\infty} f_X(x)\,dx$$

$$= 0 + \lambda \int_{0}^{\infty} e^{-\lambda x}$$

$$= \lambda \cdot \frac{1}{\lambda} = \underline{1}$$

Finding the distribution function,

$$F_X(x) = \int_{-\infty}^{x} f_X(x) = \int_{-\infty}^{0} f_X(x)\,dx + \int_{0}^{x} f_X(x)\,dx$$

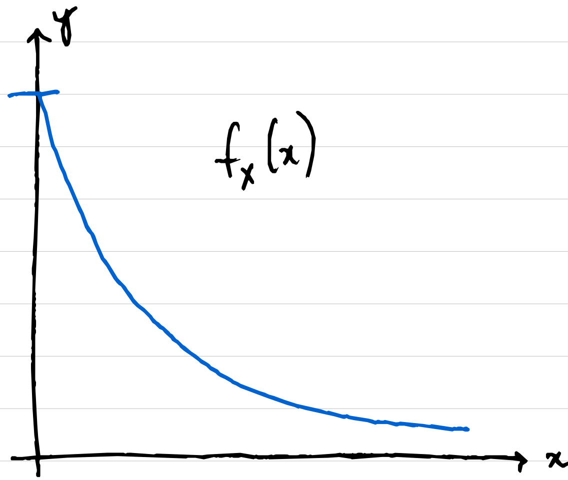
$$= 0 + \lambda \int_{0}^{x} e^{-\lambda x}\, dx$$

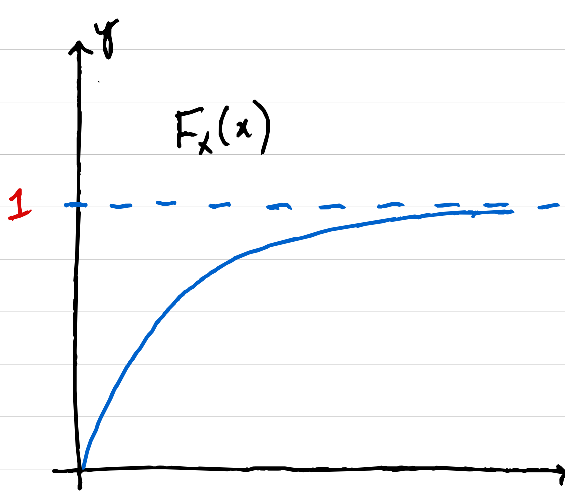
$$= \frac{\lambda}{\lambda}\left[-e^{-\lambda x}\right]_{0}^{x}$$

$$= 1 - e^{-\lambda x}$$

Thus the complete specification of distribution function is

$$\boxed{F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}}$$





asymptote
$y = 1$

$$\lim_{x \to \infty} F_X(x) = \lim_{x \to \infty} 1 - e^{-\lambda x}$$

$$= 1 - 0$$

$$= 1$$

The exponential distribution is often used to model waiting times between certain events such as natural disasters, machine breakdowns, or customers joining a queue.

If the waiting times are independant, and $\exp(\lambda)$ distributed, it can be shown that the numbers of arrivals follows $\text{Pois}(\lambda t)$ distribution.

Conversely, if numbers of arrivals is $\text{Pois}(\lambda t)$ distributed, then waiting times follow $\text{Exp}(\lambda)$ distribution.

Theorem: (Memoryless property): (for exponential distribution)
If waiting times are exponentially distributed, the probability that something doesn't occur in the next $t+s$ units of time given we've already waited s units long is the probability of something not occuring in the next t sec.

i.e

$$P(X > s+t \mid X > s) = P(X > t)$$

proof

$$P(X > s+t \mid X > s) = \frac{P(\{X > s+t\} \cap \{X > s\})}{P(X > s)}$$

$$= \frac{P(X > s+t)}{P(X > s)}$$

$$= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}$$

$$= e^{-\lambda t}$$

$$= 1 - (1 - e^{-\lambda t})$$

$$= 1 - P(X \leq t)$$

$$= P(X > t)$$

$\blacksquare$

Defn 5.7: Pareto Distribution:

We say that the continuous random variable X has the pareto distribution with parameter $\alpha > 0$ and write

$$\boxed{X \sim Par(\alpha)}$$

if the density function of X is

$$\boxed{f_X(x) = \begin{cases} \dfrac{\alpha}{x^{\alpha+1}} & \text{if } x \geq 1 \\ 0 & \text{if } x < 1 \end{cases}}$$

$f_X(x)$ is indeed a density function since

$f_X(x) \geq 0 \quad \forall x \in \mathbb{R}$ and

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_{-\infty}^{1} f_X(x)\,dx + \int_{1}^{\infty} f_X(x)\,dx$$

$$= 0 + \int_{1}^{\infty} \frac{\alpha}{x^{\alpha+1}}\,dx$$

$$= \alpha \int_1^{\infty} \frac{1}{x^{\alpha+1}} \, dx$$

$$= \alpha \left[-\frac{1}{\alpha} x^{-\alpha} \right]_1^{\infty}$$

$$= \left[-0 - (-1) \right] = 1$$

The distribution function is given by

$$F_X(x) = \int_{-\infty}^{x} f_X(x) \, dx$$

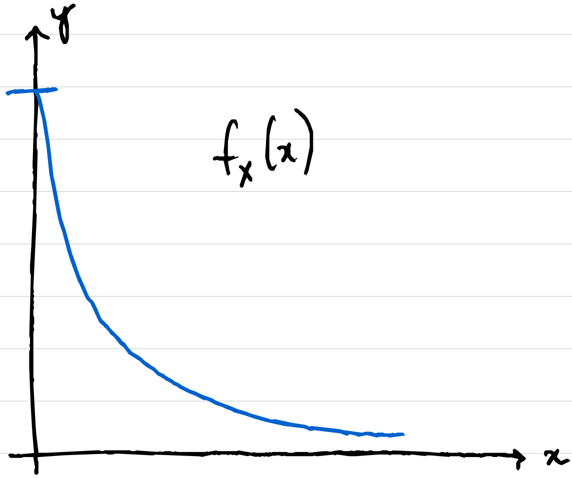
$$= \int_{-\infty}^{1} f_X(x) \, dx + \int_1^{x} f_X(x) \, dx$$

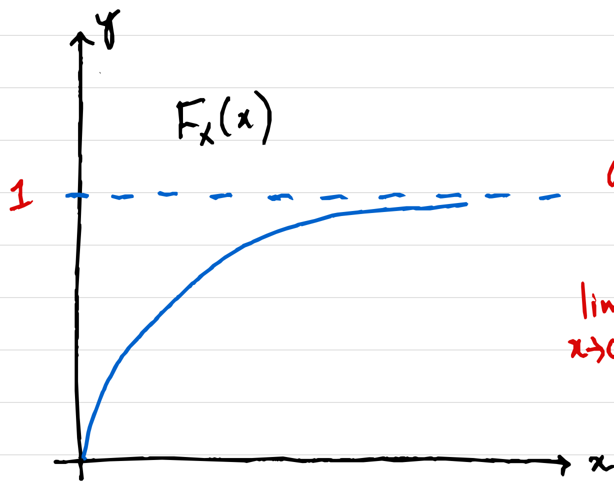
$$= 0 + \alpha \int_1^{x} \frac{1}{x^{\alpha+1}} \, dx$$

$$= \alpha \left[-\frac{1}{\alpha} x^{-\alpha} \right]_1^{x} = 1 - x^{-\alpha}$$

Thus full specification of distribution fn is

$$\boxed{F_X(x) = \begin{cases} 0 & x < 1 \\ 1 - x^{-\alpha} & x \geq 1 \end{cases}}$$





asymptote
$y = 1$

$$\lim_{x\to\infty} F_X(x) = \lim_{x\to\infty} 1 - x^{-\alpha}$$

$$= 1 - 0$$

$$= 1$$

Pareto distribution is an example of a "power law" distribution because density is falling of a power of x rather than exponentially for large x.

↳ Thus distribution has long tail.

Pareto distributions show up in many complex systems, in particular social systems.

Examples of pareto:

- Number of friends on a social network.

- Wealth and income

Defn 5.8: The Normal Distribution:

We say that the continuous random variable X has the normal distribution with mean μ and variance $\sigma^2 > 0$ and write

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

if the density function of X is

$$\boxed{f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad \text{for } x \in \mathbb{R}}$$

Note: $\exp(x)$ means e^x

If $X \sim N(\mu, \sigma^2)$ then the distribution fn is given by

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx$$

Unfortunately there is no explicit formula for F_X since f_X has no antiderivative

However as we shall see in chapter 8, any $X \sim N(\mu, \sigma^2)$ can be converted to $X \sim N(0,1)$ by a transformation.

↳ as a result a simple table suffices in calculating any $X \sim N(\mu, \sigma^2)$

Because $X \sim N(0,1)$ is used so often, standard symbols have been introduced.

$X \sim N(0,1)$ has density function:

$$\boxed{f_X(x) = \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}}$$

distribution function denoted by Φ

$$\boxed{F_X(x) = \Phi(x) = \int_{-\infty}^{x} \phi(s)\, ds}$$

Note: <u>properties of $X \sim N(0,1)$</u>

- $\phi(x)$ is symmetric about zero:

$$\phi(-x) = \phi(x)$$

- The table for Φ do not contain values for $\Phi(a)$ but for "right tail probabilities"


– right tail

$$P(X \geq a) = 1 - \Phi(a)$$

So to calculate $\Phi(a)$ (left tail) →


left tail

$$\Phi(a) = 1 - P(X \geq a) = P(X \leq a)$$

$\uparrow$ given in table right tail.

5.3 Quantiles

Defn 5.9: Let X be a random variable with distribution function F_X and let $p \in [0,1]$

The p^{th} quantile or $(100 \cdot p)$th percentile of the distribution X is the smallest number q_p such that

$$F_X(q_p) = P(X \leq q_p) = p$$

For example if

$$P(X \leq q_p) = 0.1,$$

q_p is called the 0.1th quantile or 10th percentile.

- The median is 50th percentile
- Upper quartile is 75^{th} percentile
- Lower quartile is 25^{th} percentile

Example: 5.10 Let $X \sim Exp(\lambda)$
Calculate Median of X.

Solution: Let q denote median of X.
So q satisfies

$$F_X(q) = P(X \le q) = 0.5$$

$$\Rightarrow 1 - e^{-\lambda q} = 1/2$$

$$\Rightarrow \boxed{q = \frac{\log(2)}{\lambda}}$$

$\square$

In general, for continuous random variables, q_p is often easy to find.

F is strictly increasing from 0 to 1 by Thm 4.8
so

$$\boxed{q_p = F_X^{inv}(x)} \quad \text{(inverse of } F_X\text{)}$$

or

$$\boxed{q_p = F_X^{-1}(p)}$$

Use this formula for continuous random variables.