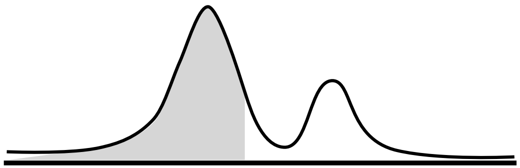
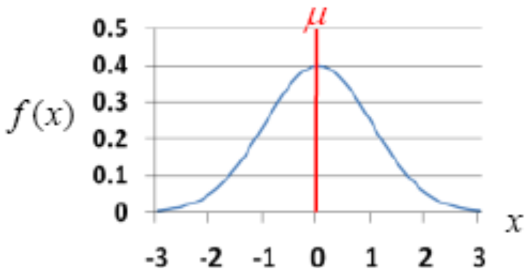
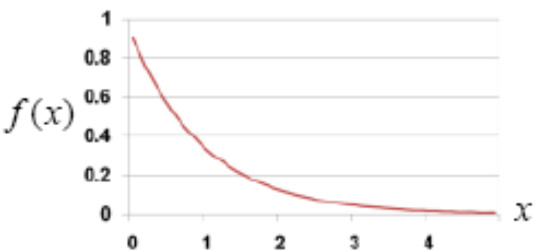
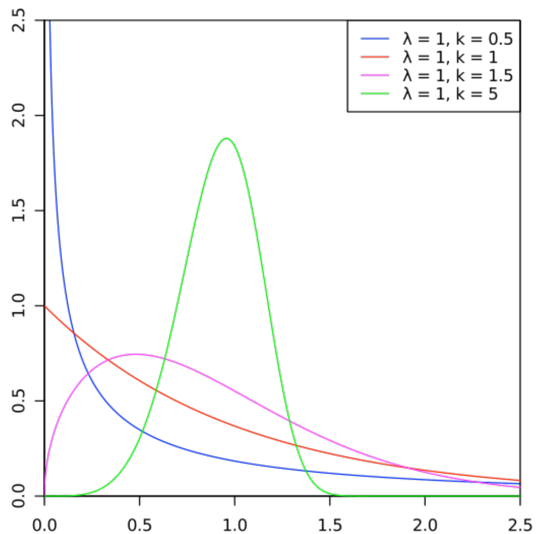




continuous random variables





Discrete random variable: takes values in a finite or countable set, e.g.

$X \in \{1,2, \ldots, 6\}$ with equal probability

X is positive integer i with probability 2^{-i}

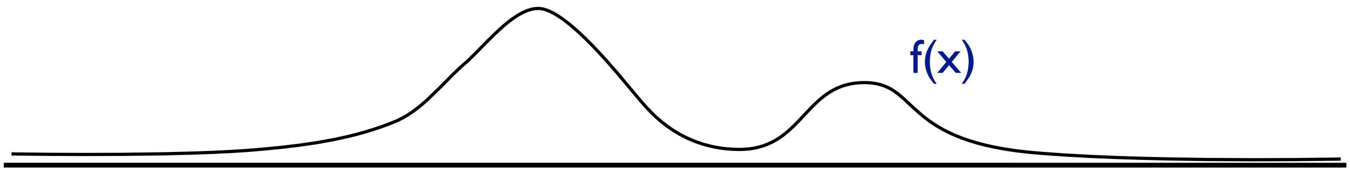
Continuous random variable: takes values in an uncountable set, e.g.

X is the weight of a random person (a real number)

X is a randomly selected point inside a unit square

X is the waiting time until the next packet arrives at the server

$f(x): R \rightarrow R$, the *probability density function* (or simply "density")



Require:

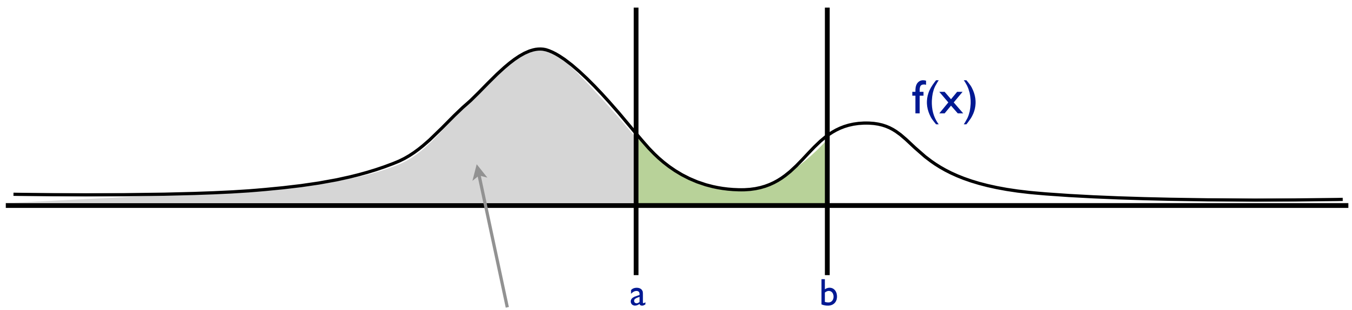
I.e., distribution is:

$f(x) \geq 0$, and ← nonnegative, and

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

just like discrete PMF

F(x): the *cumulative distribution function* (aka the "distribution")

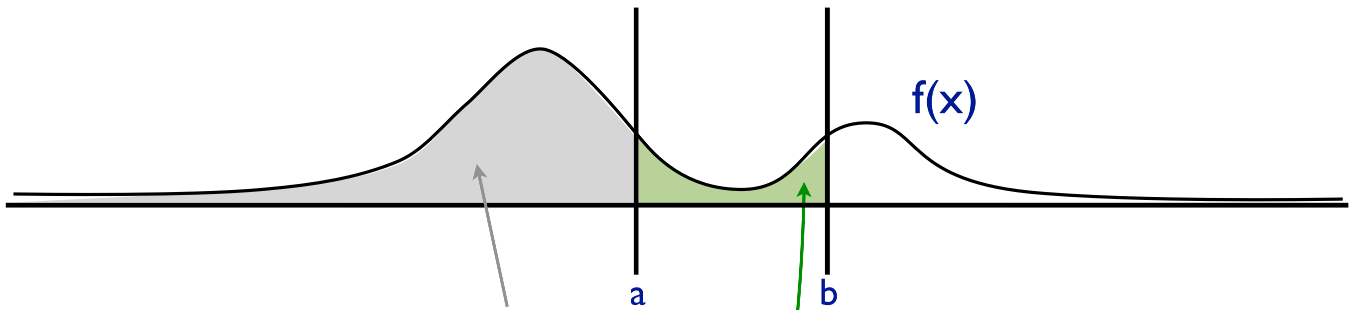


$F(a) = P(X \le a) = \int_{-\infty}^{a} f(x)\,dx$ (Area left of a)

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

cdf

F(x): the *cumulative distribution function* (aka the "distribution")

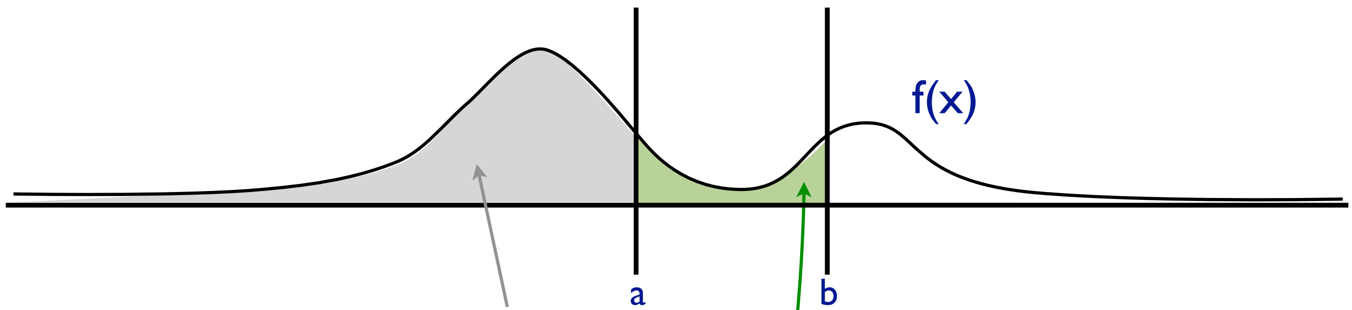


$F(a) = P(X \leq a) \;\; = \int_{-\infty}^{a} f(x)\, dx$ (Area left of a)

$P(a < X \leq b) = F(b) - F(a)$ (Area between a and b)

cdf

F(x): the *cumulative distribution function* (aka the "distribution")

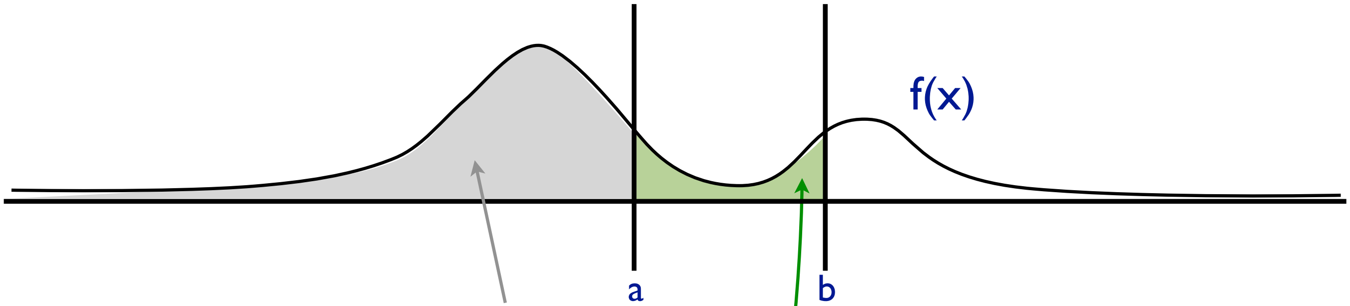


$F(a) = P(X \leq a)$ $= \int_{-\infty}^{a} f(x)\, dx$ (Area left of a)

$P(a < X \leq b) = F(b) - F(a)$ (Area between a and b)

Relationship between f(x) and F(x)?

F(x): the *cumulative distribution function* (aka the "distribution")



$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)\,dx$ (Area left of a)

$P(a < X \leq b) = F(b) - F(a)$ (Area between a and b)

A key relationship:

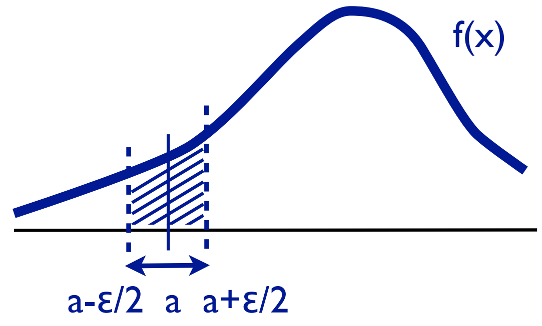
$f(x) = \frac{d}{dx} F(x)$, since $F(a) = \int_{-\infty}^{a} f(x)\,dx$, Densities are *not* probabilities; e.g. may be > 1

$P(X = a) = \lim_{\varepsilon \to 0} P(a-\varepsilon < X \leq a) = F(a)-F(a) = 0$

I.e., the probability that a continuous r.v. falls *at* a specified point is *zero*.

But

the probability that it falls *near* that point is *proportional to the density*:



why is it called a density?

Densities are *not* probabilities; e.g. may be > 1

$P(X = a) = \lim_{\varepsilon \to 0} P(a-\varepsilon < X \leq a) = F(a)-F(a) = 0$

I.e.,

the probability that a continuous r.v. falls *at* a specified point is *zero.*

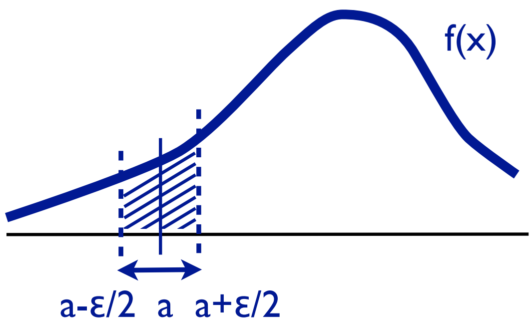
But

the probability that it falls *near* that point is *proportional to the density*:

$P(a - \varepsilon/2 < X \leq a + \varepsilon/2) =$

$F(a + \varepsilon/2) - F(a - \varepsilon/2)$

$\approx \varepsilon \cdot f(a)$



I.e., in a large random sample, expect more samples where density is higher (hence the name "density").

Much of what we did with discrete r.v.s carries over almost unchanged, with Σ_x... replaced by $\int$... dx

E.g.

For discrete r.v. X, $E[X] = \Sigma_x \, xp(x)$

For continuous r.v. X, $E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \, dx$

Much of what we did with discrete r.v.s carries over almost unchanged, with $\Sigma_x \ldots$ replaced by $\int \ldots \, dx$

E.g.

For discrete r.v. X, $E[X] = \Sigma_x \, xp(x)$

For continuous r.v. X, $E[X] = \int_{-\infty}^{\infty} x \cdot f(x) \, dx$

Why?

(a) We define it that way

(b) The probability that X falls "near" x, say within x±dx/2, is ≈f(x)dx, so the "average" X should be ≈ Σ xf(x)dx (summed over grid points spaced dx apart on the real line) and the limit of that as dx→0 is ∫xf(x)dx

Continuous random variable X has density f(x), and

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

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

$$E[X^2] = \int_{-\infty}^{\infty} x^2 \cdot f(x)\,dx$$

Linearity

$E[aX+b] = aE[X]+b$

$E[X+Y] = E[X]+E[Y]$

still true, just as for discrete

Functions of a random variable

$E[g(X)] = \int g(x)f(x)dx$

just as for discrete, but w/integral

Alternatively, let $Y = g(X)$, find the density of Y, say f_Y, and directly compute $E[Y] = \int y f_Y(y) dy$.

Definition is same as in the discrete case

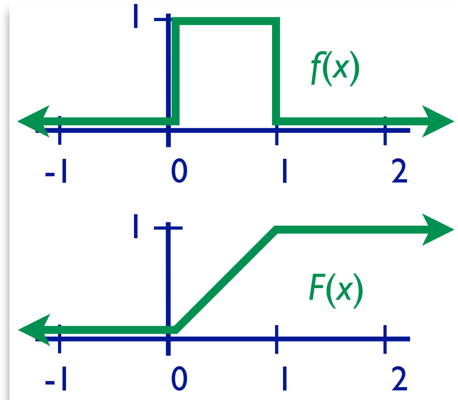
$\mathrm{Var}[X] = E[(X-\mu)^2]$ where $\mu = E[X]$

Identity still holds:

$\mathrm{Var}[X] = E[X^2] - (E[X])^2$

proof "same"

Let
$$f(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$



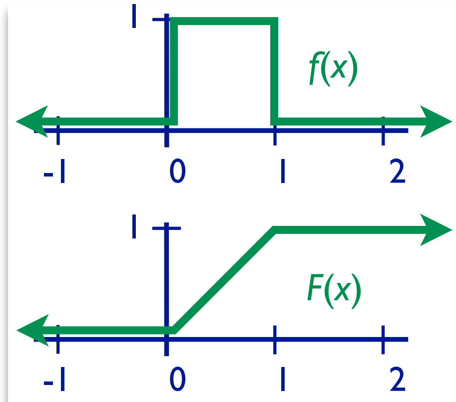
What is F(x)? What is E(X)?

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

$$E(X) = \int_0^1 x f(x)\,dx = \int_0^1 x\,dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}$$

$$E(X^2) = \int_0^1 x^2 f(x)\,dx = \frac{1}{3}$$

Let $f(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$



$$F(a) = \int_{-\infty}^{a} f(x)dx$$

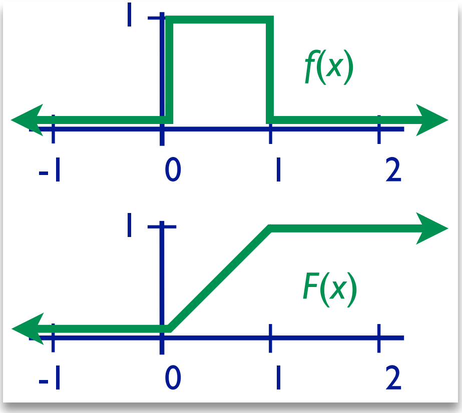
$$= \begin{cases} 0 & \text{if } a \le 0 \\ a & \text{if } 0 < a \le 1 \text{ (since } a = \int_0^a 1dx) \\ 1 & \text{if } 1 < a \end{cases}$$

$$\underline{E[X]} = \int_{-\infty}^{\infty} xf(x)dx = \int_0^1 x\,dx = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2}$$

$$\underline{E[X^2]} = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^1 x^2\,dx = \left.\frac{x^3}{3}\right|_0^1 = \frac{1}{3}$$

$Var(X) = \frac{1}{3} - \left(\frac{1}{2}\right)^2$

Let $f(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$



$$F(a) = \int_{-\infty}^{a} f(x)dx$$

$$= \begin{cases} 0 & \text{if } a \leq 0 \\ a & \text{if } 0 < a \leq 1 \text{ (since } a = \int_0^a 1dx) \\ 1 & \text{if } 1 < a \end{cases}$$

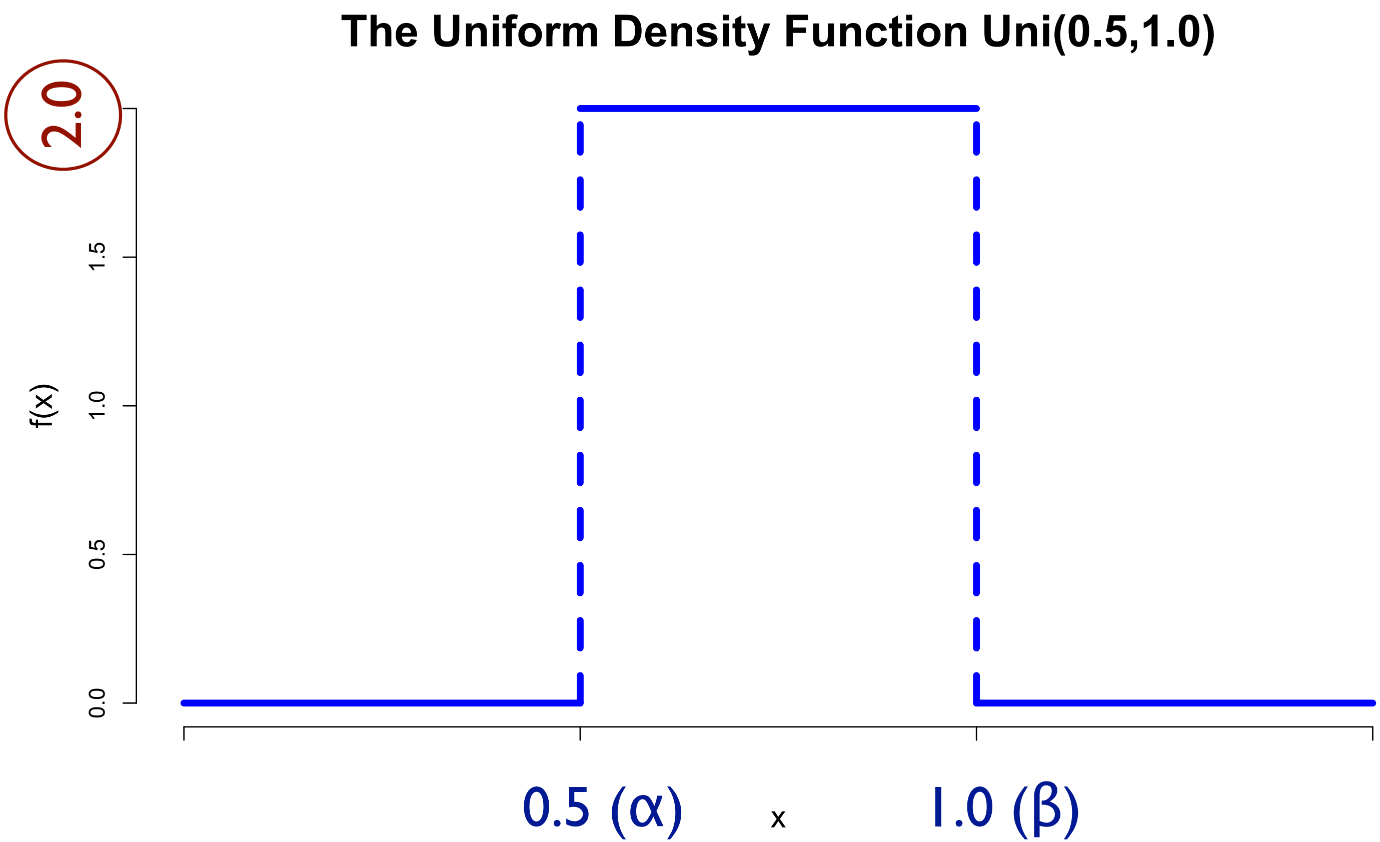
$$E[X] = \int_{-\infty}^{\infty} x f(x)dx = \int_0^1 x\,dx = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2}$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)dx = \int_0^1 x^2\,dx = \left.\frac{x^3}{3}\right|_0^1 = \frac{1}{3}$$

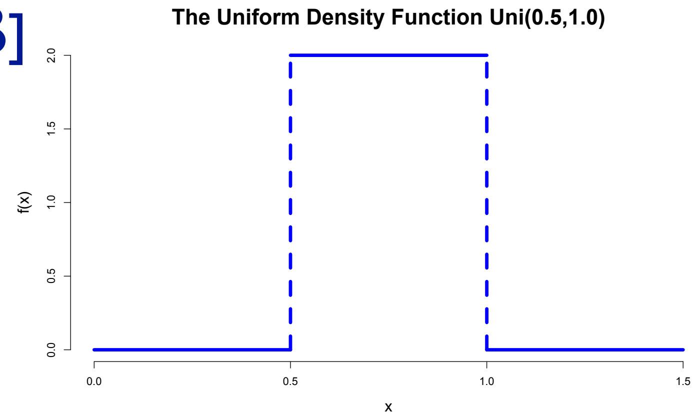
$$\text{Var}[X] = E[X^2] - (E[X])^2 = \tfrac{1}{3} - \tfrac{1}{4} = \tfrac{1}{12} \quad (\sigma \approx 0.29)$$

X ~ Uni(α,β) is uniform in [α,β]

$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha, \beta] \\ 0 & \text{otherwise} \end{cases}$$



X ~ Uni(α,β) is uniform in [α,β]



$$f(x) = \begin{cases} \frac{1}{\beta-\alpha} & x \in [\alpha, \beta] \\ 0 & \text{otherwise} \end{cases}$$

$$\Pr(a \leq X \leq b) = \int_a^b f(x)\,dx = \frac{b-a}{\beta-\alpha}$$

if α≤a≤b≤β:

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx = \frac{\alpha+\beta}{2}$$

Yes, you should review your basic calculus; e.g., these 2 integrals would be good practice.

Radioactive decay: How long until the next alpha particle?

Customers: how long until the next customer/packet arrives at the checkout stand/server?

Buses: How long until the next #71 bus arrives on the Ave?

Yes, they have a schedule, but given the vagaries of traffic, riders with-bikes-and-baby-carriages, etc., can they stick to it?

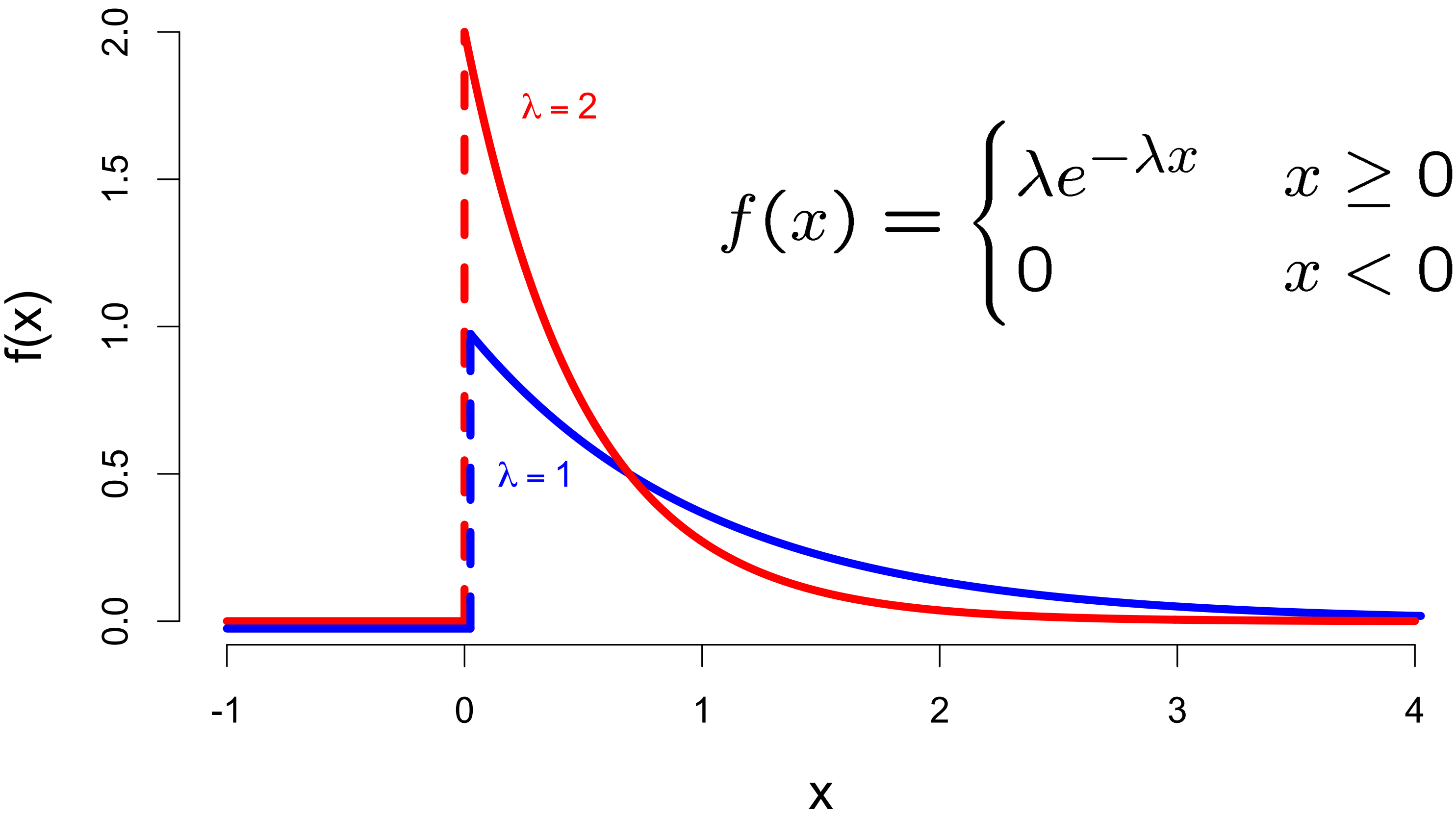
Assuming events are independent, happening at some fixed *average* rate of λ per unit time – the waiting time until the next event is exponentially distributed (next slide)

interval dt of time Prob event $\lambda\, dt$

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

Exponential with param λ

The Exponential Density Function



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

X ~ Exp(λ)

$E(X) = \int_0^\infty x \lambda e^{-\lambda x} \, dx = \frac{1}{\lambda}$

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

$$E[X] = \frac{1}{\lambda} \qquad \mathrm{Var}[X] = \frac{1}{\lambda^2}$$

$$\Pr(X \geq t) = e^{-\lambda t} = 1\text{-}F(t) = \int_t^\infty \lambda e^{-\lambda x} dx = -e^{-\lambda x}\Big|_t^\infty$$

Memorylessness:

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

Assuming exp distr, if you've waited *s* minutes, prob of waiting *t* more is exactly same as *s = 0*

Relation to Poisson

Same process, different measures:

Poisson: *how many* events in a *fixed time*;

Exponential: *how long* until the *next event*

λ is avg # per unit time;

$1/\lambda$ is mean wait