
Anna Karlin
Most Slides by Alex Tsun

Poisson RV Example

Suppose Lookbook gets on average 120 new users per hour, and Quickgram gets 180 new users per hour, independently. What is the probability that, combined, less than 2 users sign up in the next minute?

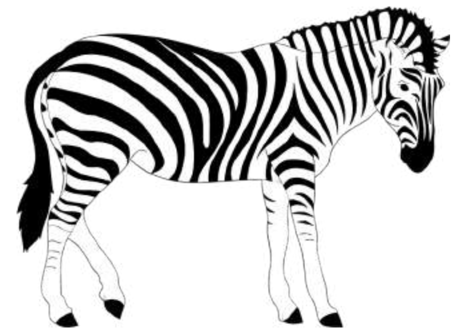
Convert λ's to the same unit of interest. For us, it's a minute.

$$X \sim Poi(2 \text{ users/min}) \qquad Y \sim Poi(3 \text{ users/min})$$

$$Z = X + Y \sim Poi(2+3) = Poi(5)$$

$$P(Z < 2) = p_Z(0) + p_Z(1) = e^{-5}\frac{5^0}{0!} + e^{-5}\frac{5^1}{1!} = 6e^{-5} \approx 0.04$$

The Zoo of Discrete RV's



- The Bernoulli RV
- The Binomial RV
- The Geometric RV
- The Uniform RV
- The Poisson RV
- The Negative Binomial RV
- The Hypergeometric RV

Important Examples:

Uniform(a,b): $P(X = i) = \frac{1}{b - a + 1}$ $\mu = \frac{a+b}{2}, \sigma^2 = \frac{(b-a)(b-a+2)}{12}$

Bernoulli(p): $P(X = 1) = p$, $P(X = 0) = 1\text{-}p$ $\mu = p$, $\sigma^2 = p(1\text{-}p)$

Binomial(n,p) $P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$ $\mu = np$, $\sigma^2 = np(1\text{-}p)$

Poisson(λ): $P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}$ $\mu = \lambda$, $\sigma^2 = \lambda$

Bin(n,p) ≈ Poi(λ) where $\lambda = np$ fixed, $n \to \infty$ (and so $p = \lambda/n \to 0$)

Geometric(p) $P(X = k) = (1\text{-}p)^{k-1}p$ $\mu = 1/p$, $\sigma^2 = (1\text{-}p)/p^2$

Probability

4.1 Continuous random Variables Basics

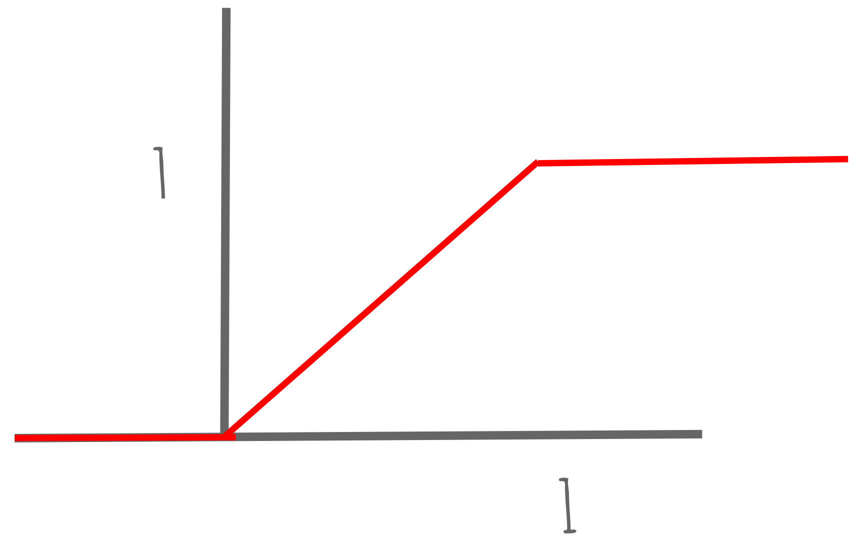
Anna Karlin
Most Slides by Alex Tsun

Agenda

- Probability Density Functions (PDFs)
- Cumulative Distribution Functions (CDFs)
- From Discrete to Continuous

CDF Intuition

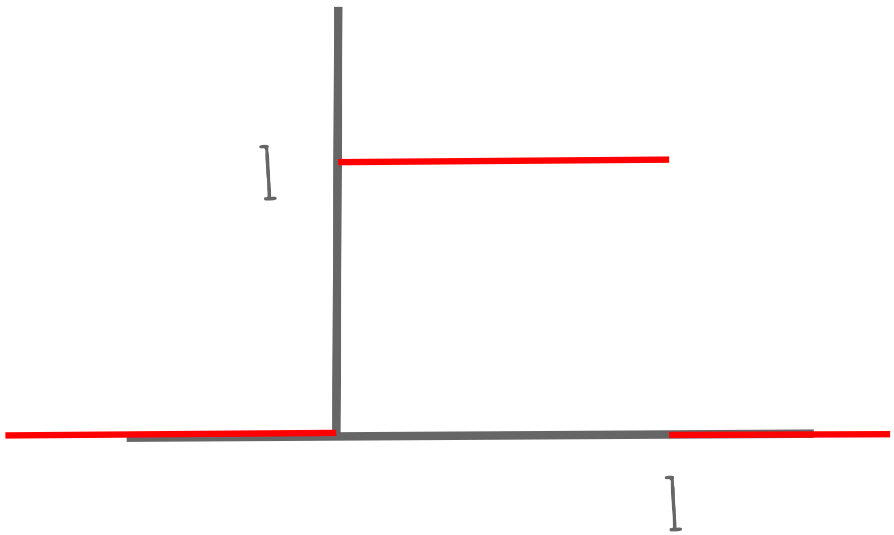
$$F_X(w) = P(X \leq w)$$



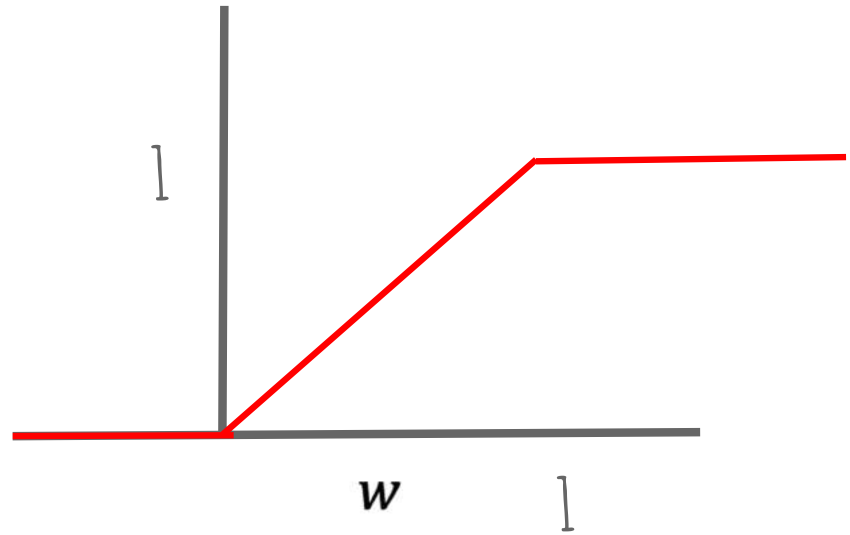
$$F_X(w) = \begin{cases} 0, & w < 0 \\ w, & 0 \leq w \leq 1 \\ 1, & w > 1 \end{cases}$$

CDF Intuition

$f_X(v)$



$F_X(w) = P(X \leq w)$

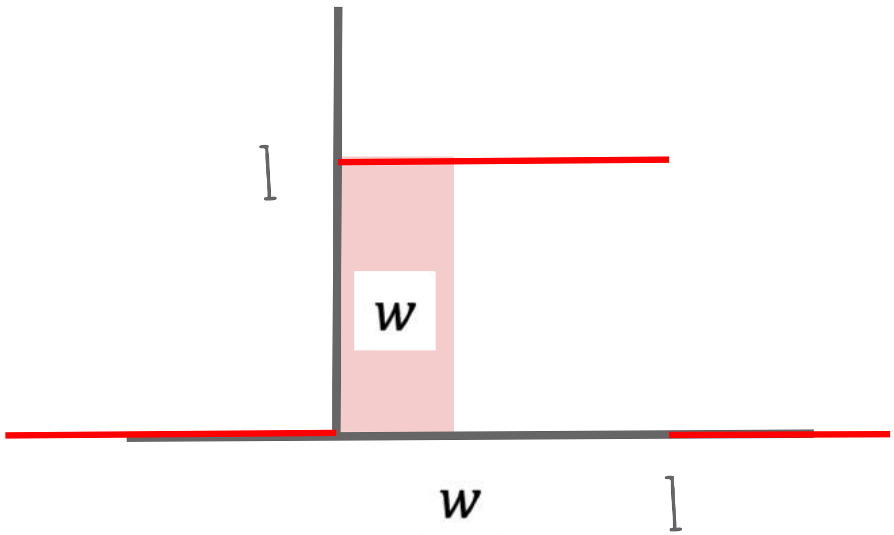


$$f_X(v) = \begin{cases} 1, & 0 \leq v \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

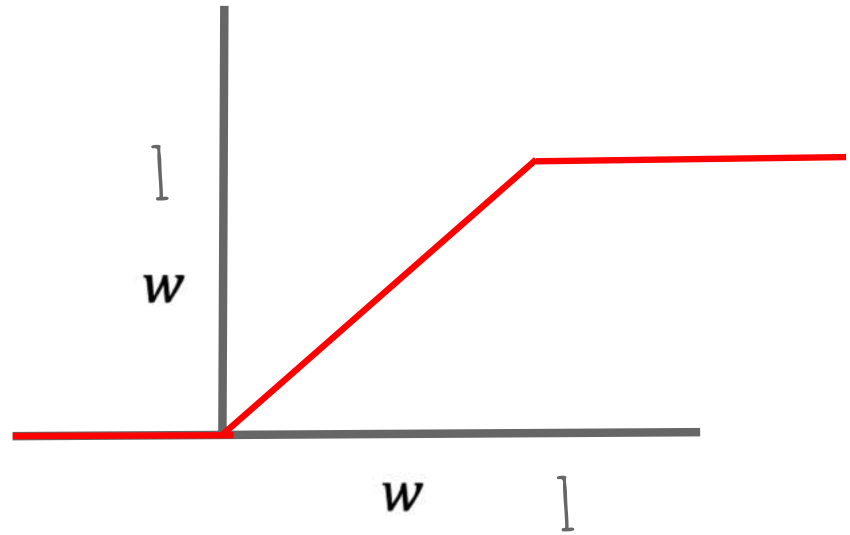
$$F_X(w) = \begin{cases} 0, & w < 0 \\ w, & 0 \leq w \leq 1 \\ 1, & w > 1 \end{cases}$$

CDF Intuition

$f_X(v)$



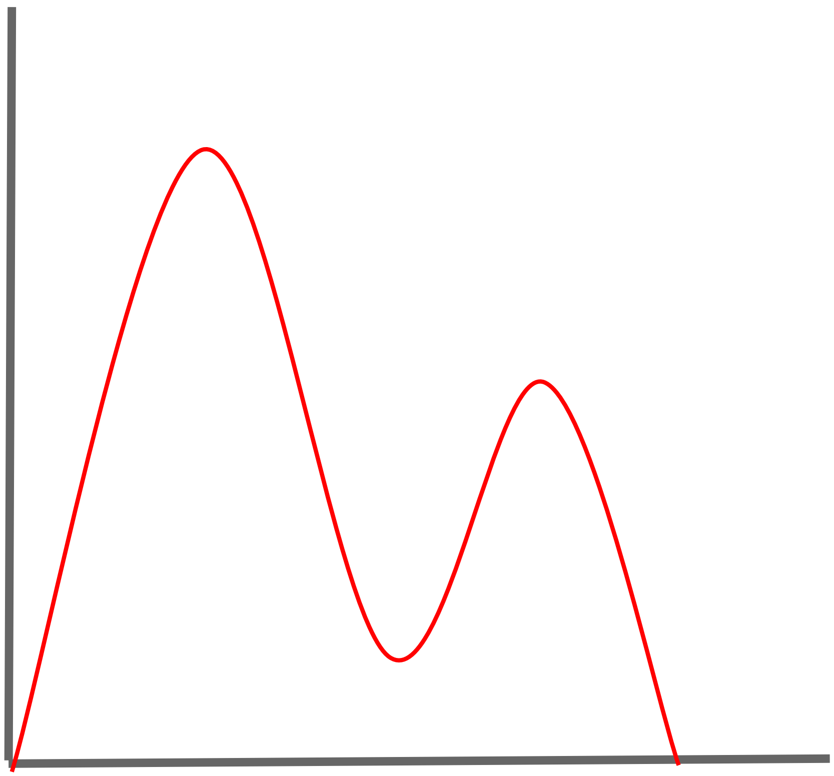
$F_X(w) = P(X \leq w)$



$$f_X(v) = \begin{cases} 1, & 0 \leq v \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

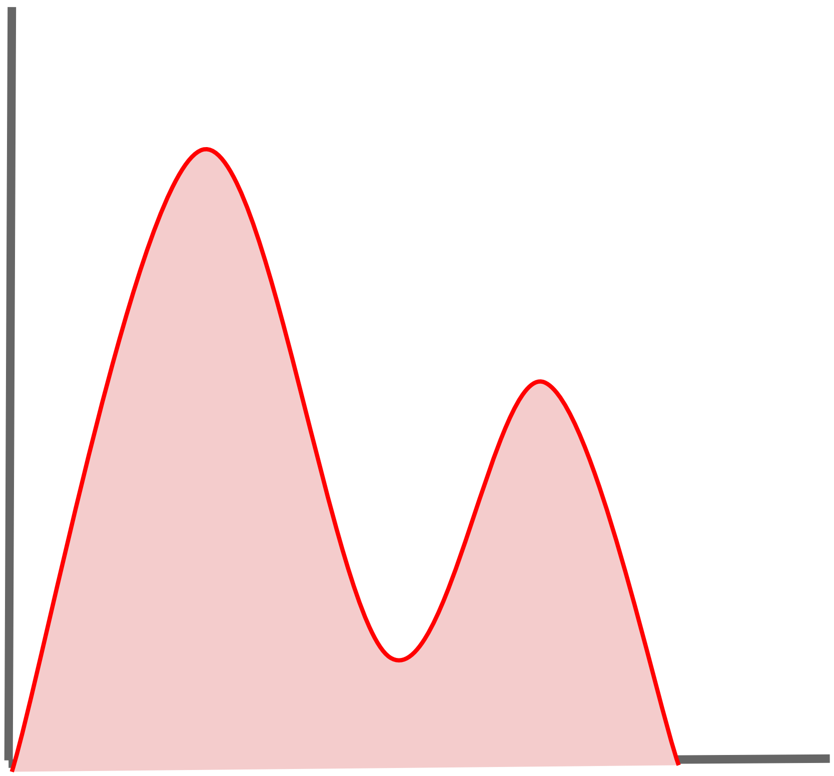
$$F_X(w) = \begin{cases} 0, & w < 0 \\ w, & 0 \leq w \leq 1 \\ 1, & w > 1 \end{cases}$$

PDF Intuition



$f_X(z) \geq 0$ for all $z \in \mathbb{R}$

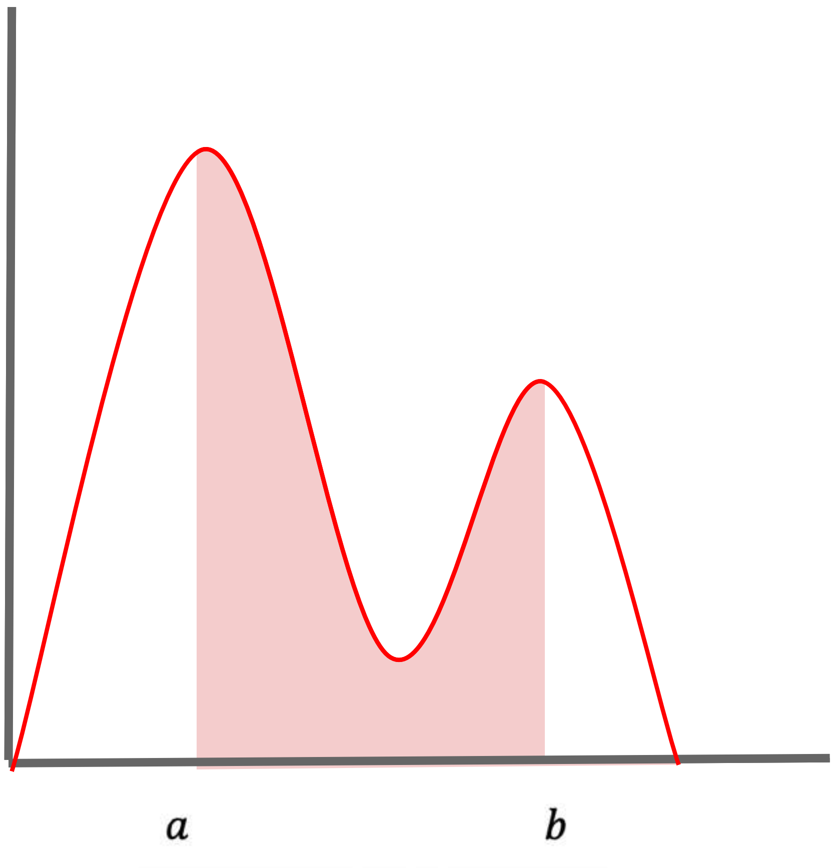
PDF Intuition



$$f_X(z) \geq 0 \text{ for all } z \in \mathbb{R}$$

$$\int_{-\infty}^{\infty} f_X(t)dt = 1$$

PDF Intuition

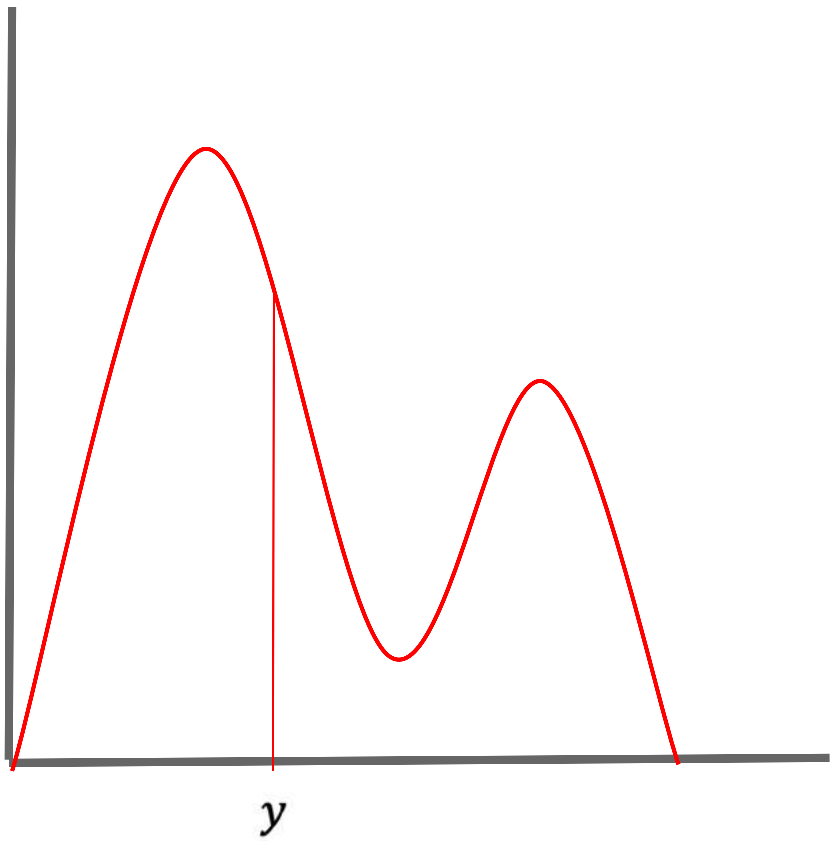


$$f_X(z) \geq 0 \text{ for all } z \in \mathbb{R}$$

$$\int_{-\infty}^{\infty} f_X(t)\,dt = 1$$

$$P(a \leq X \leq b) = \int_a^b f_X(w)\,dw$$

PDF Intuition



$$f_X(z) \geq 0 \text{ for all } z \in \mathbb{R}$$

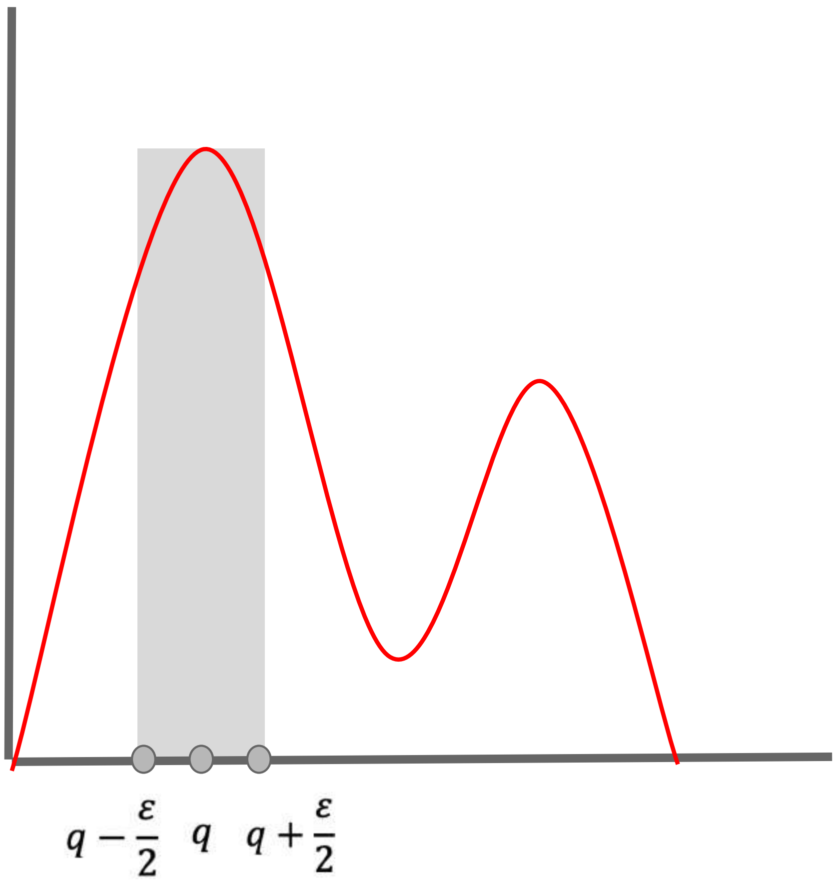
$$\int_{-\infty}^{\infty} f_X(t)dt = 1$$

$$P(a \leq X \leq b) = \int_a^b f_X(w)dw$$

$$P(X = y) = P(y \leq X \leq y) = \int_y^y f_X(w)dw = 0$$

PDF Intuition

PDF



$$f_X(z) \geq 0 \text{ for all } z \in \mathbb{R}$$

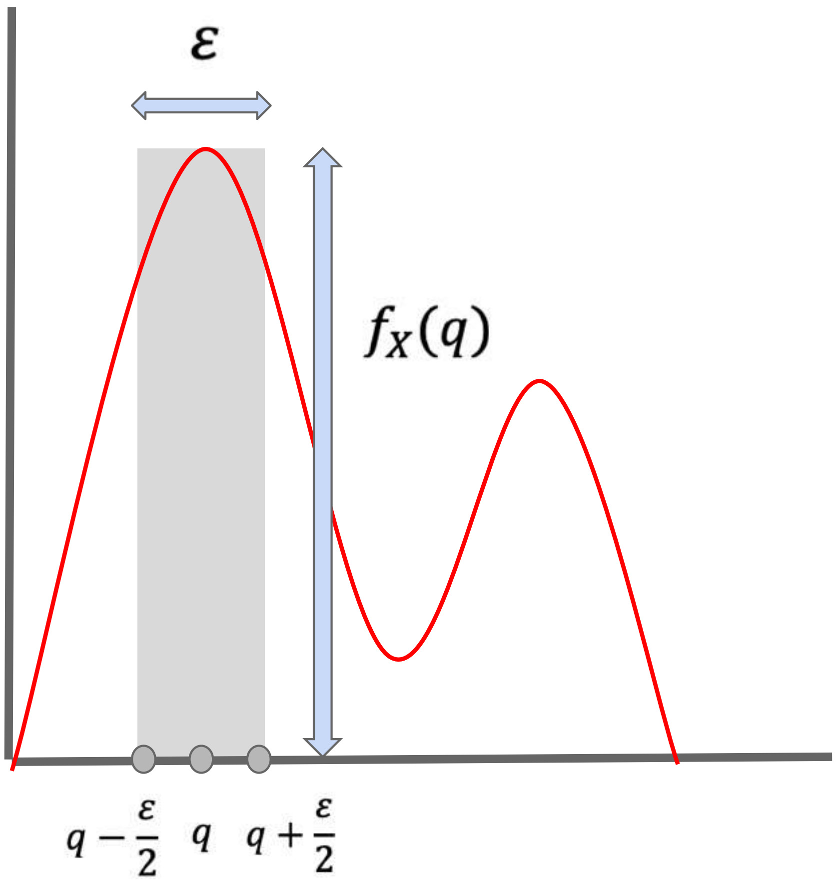
$$\int_{-\infty}^{\infty} f_X(t)dt = 1$$

$$P(a \leq X \leq b) = \int_a^b f_X(w)dw$$

$$P(X = y) = P(y \leq X \leq y) = \int_y^y f_X(w)dw = 0$$

$$P(X \approx q) \approx P\left(q - \frac{\varepsilon}{2} \leq X \leq q + \frac{\varepsilon}{2}\right)$$

PDF Intuition



$$f_X(z) \geq 0 \text{ for all } z \in \mathbb{R}$$

$$\int_{-\infty}^{\infty} f_X(t)dt = 1$$

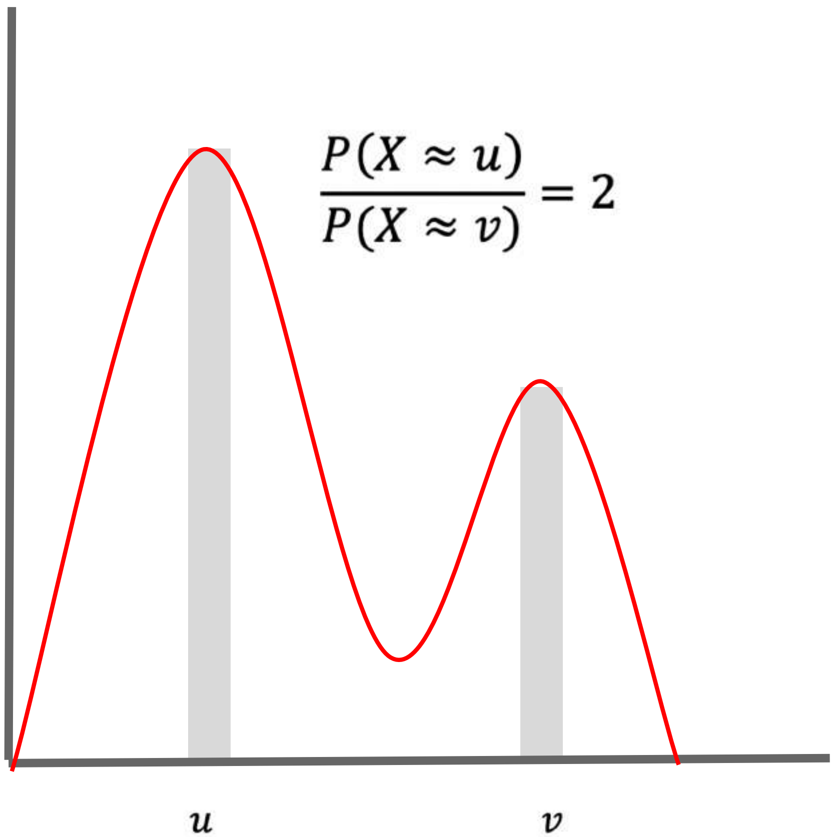
$$P(a \leq X \leq b) = \int_a^b f_X(w)dw$$

$$P(X = y) = P(y \leq X \leq y) = \int_y^y f_X(w)dw = 0$$

$$P(X \approx q) \approx P\left(q - \frac{\varepsilon}{2} \leq X \leq q + \frac{\varepsilon}{2}\right) \approx \varepsilon f_X(q)$$

PDF Intuition

$$\frac{P(X \approx u)}{P(X \approx v)} = 2$$



$$f_X(z) \geq 0 \text{ for all } z \in \mathbb{R}$$

$$\int_{-\infty}^{\infty} f_X(t)dt = 1$$

$$P(a \leq X \leq b) = \int_a^b f_X(w)dw$$

$$P(X = y) = P(y \leq X \leq y) = \int_y^y f_X(w)dw = 0$$

$$P(X \approx q) \approx P\left(q - \frac{\varepsilon}{2} \leq X \leq q + \frac{\varepsilon}{2}\right) \approx \varepsilon f_X(q)$$

$$\frac{P(X \approx u)}{P(X \approx v)} = \frac{\varepsilon f_X(u)}{\varepsilon f_X(v)} = \frac{f_X(u)}{f_X(v)}$$

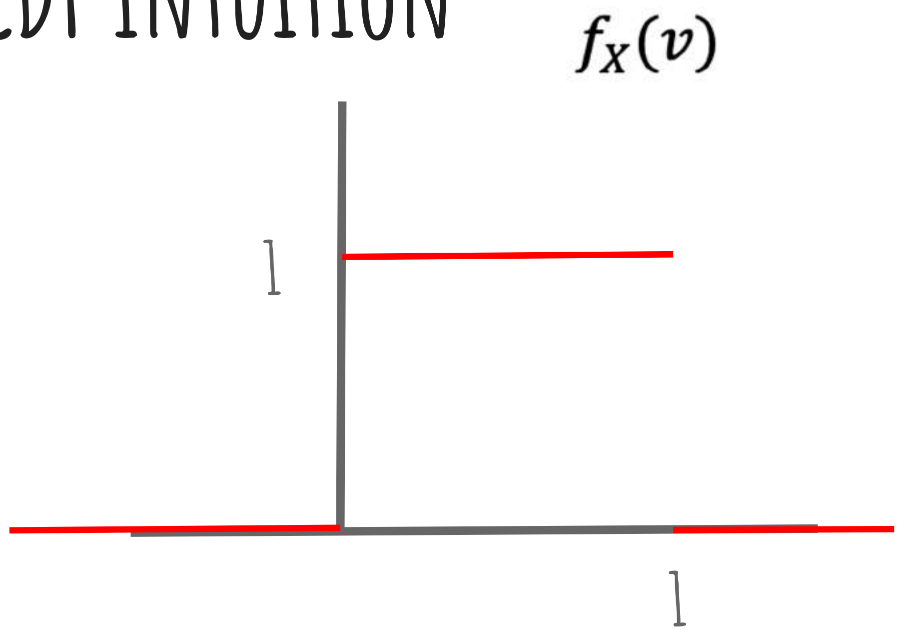
Probability Density Functions (PDFs)

Probability Density Function (PDF): Let X be a continuous rv (one whose range is typically an interval or union of intervals). The probability density function (PDF) of X is the function $f_X: \mathbb{R} \to \mathbb{R}$ such that

- $f_X(z) \geq 0$ for all $z \in \mathbb{R}$.
- $\int_{-\infty}^{\infty} f_X(t)dt = 1$.
- $P(a \leq X \leq b) = \int_a^b f_X(w)dw$.
- $P(X = y) = 0$ for any $y \in \mathbb{R}$.
- The probability that X is close to q is proportional to $f_X(q)$: $P(X \approx q) \approx P\left(q - \frac{\varepsilon}{2} \leq X \leq q + \frac{\varepsilon}{2}\right) \approx \varepsilon f_X(q)$.
- Ratios of probabilities of being near points are maintained: $\frac{P(X \approx u)}{P(X \approx v)} = \frac{\varepsilon f_X(u)}{\varepsilon f_X(v)} = \frac{f_X(u)}{f_X(v)}$.

Random Picture

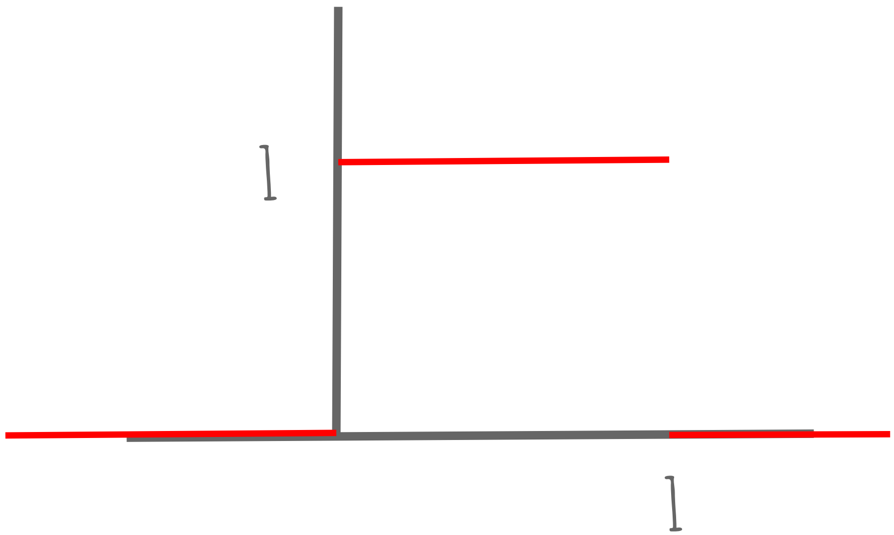
CDF Intuition



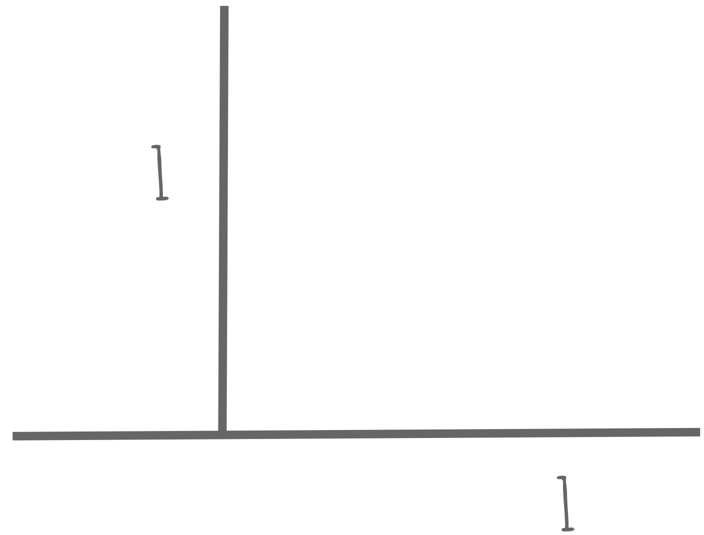
$$f_X(v) = \begin{cases} 1, & 0 \le v \le 1 \\ 0, & \text{otherwise} \end{cases}$$

CDF Intuition

$f_X(v)$



$F_X(w) = P(X \leq w)$

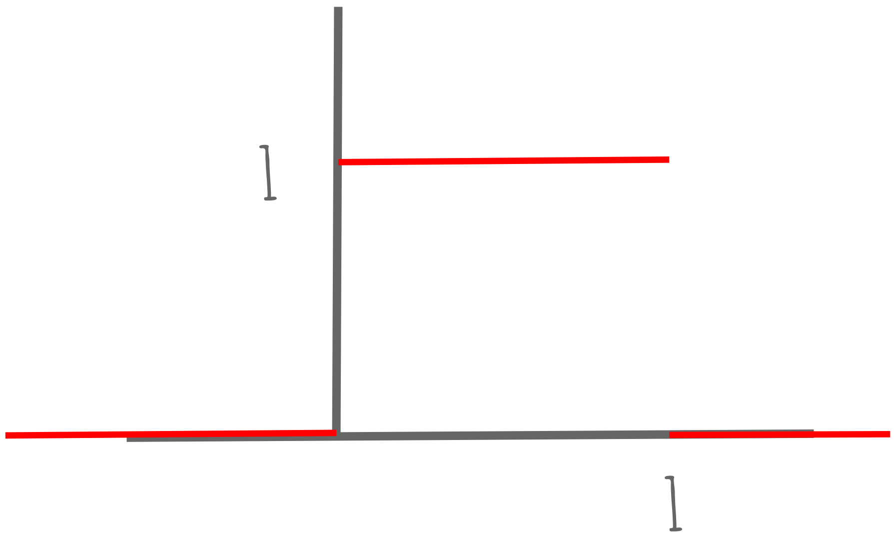


$$f_X(v) = \begin{cases} 1, & 0 \leq v \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

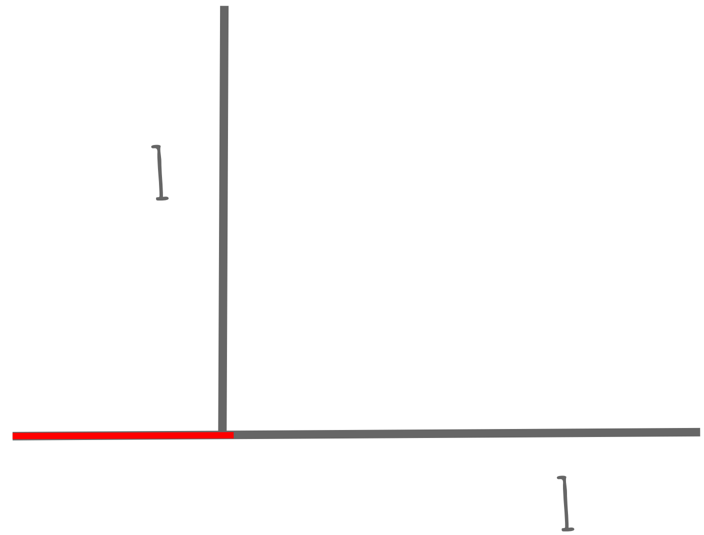
$$F_X(w) = \begin{cases} & w < 0 \\ & 0 \leq w \leq 1 \\ & w > 1 \end{cases}$$

CDF Intuition

$f_X(v)$



$F_X(w) = P(X \leq w)$

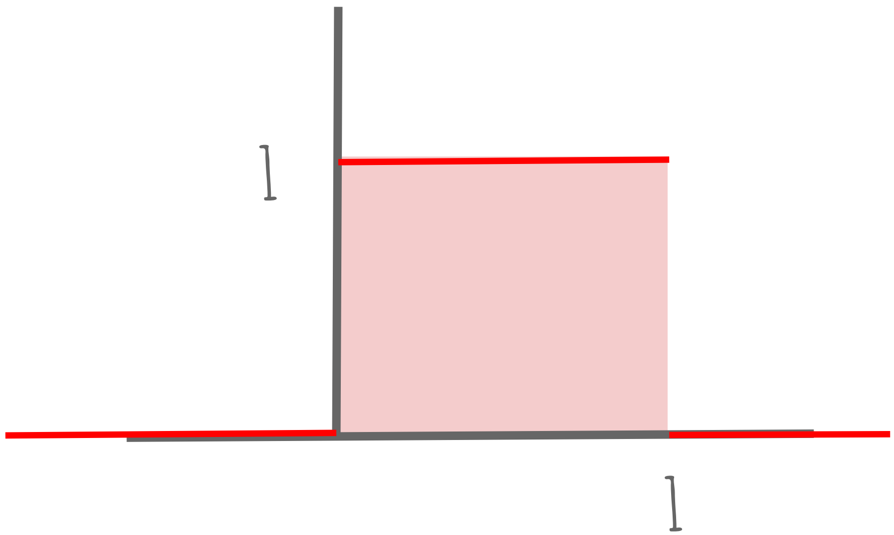


$$f_X(v) = \begin{cases} 1, & 0 \leq v \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

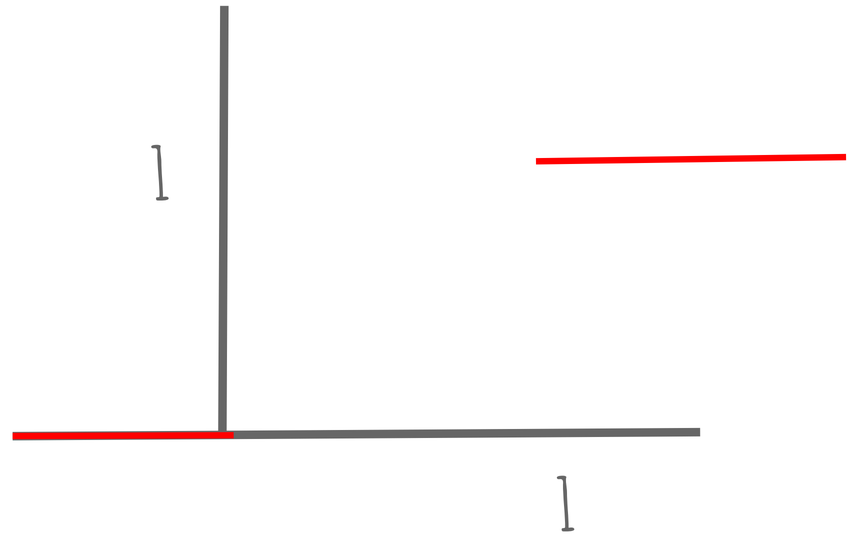
$$F_X(w) = \begin{cases} 0, & w < 0 \\ & 0 \leq w \leq 1 \\ & w > 1 \end{cases}$$

CDF Intuition

$f_X(v)$



$F_X(w) = P(X \le w)$

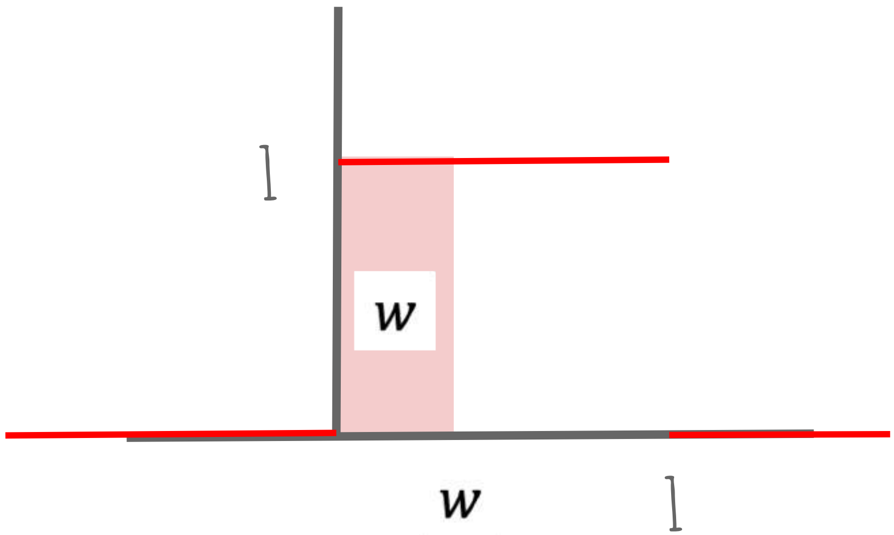


$$f_X(v) = \begin{cases} 1, & 0 \le v \le 1 \\ 0, & \text{otherwise} \end{cases}$$

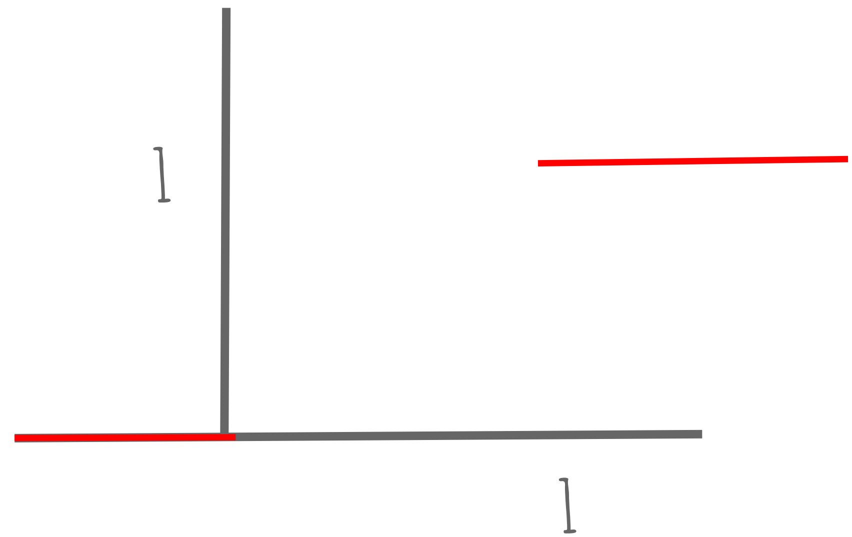
$$F_X(w) = \begin{cases} 0, & w < 0 \\ & 0 \le w \le 1 \\ 1, & w > 1 \end{cases}$$

CDF Intuition

$f_X(v)$



$F_X(w) = P(X \le w)$

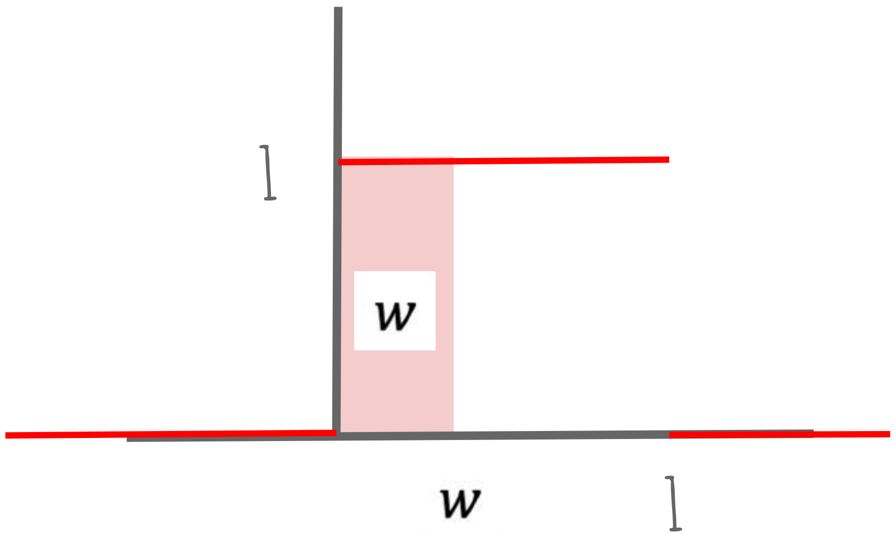


$$f_X(v) = \begin{cases} 1, & 0 \le v \le 1 \\ 0, & \text{otherwise} \end{cases}$$

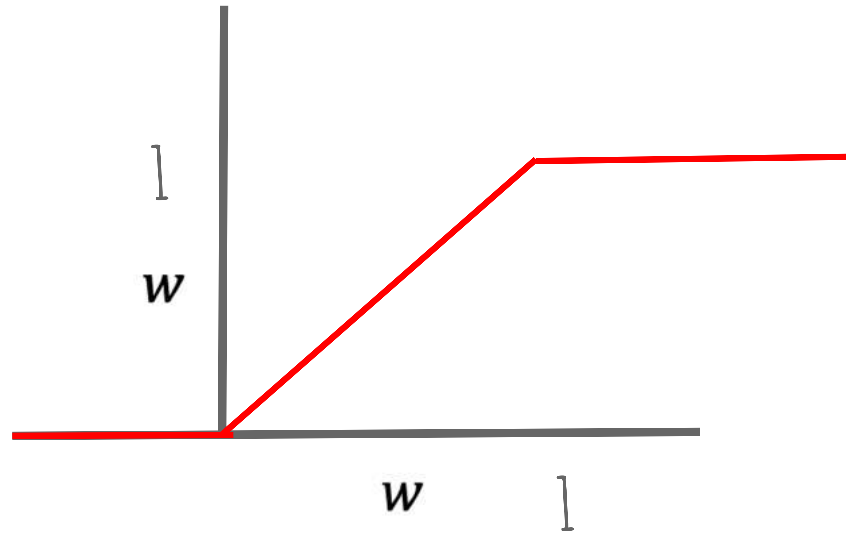
$$F_X(w) = \begin{cases} 0, & w < 0 \\ & 0 \le w \le 1 \\ 1, & w > 1 \end{cases}$$

CDF Intuition

$f_X(v)$



$F_X(w) = P(X \leq w)$



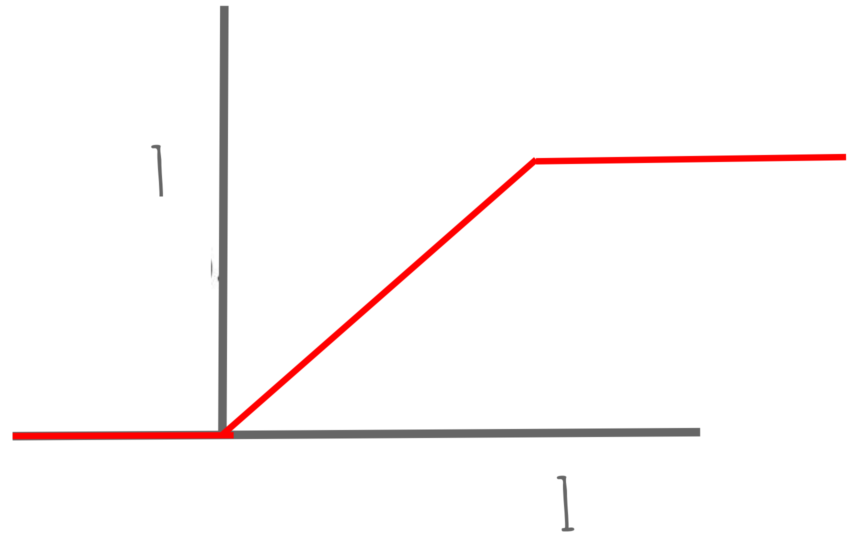
$$f_X(v) = \begin{cases} 1, & 0 \leq v \leq 1 \\ 0, & \text{otherwise} \end{cases}$$

$$F_X(w) = \begin{cases} 0, & w < 0 \\ w, & 0 \leq w \leq 1 \\ 1, & w > 1 \end{cases}$$

CDF Intuition

$F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f_X(w)dw$ for all $t \in \mathbb{R}$.

$$F_X(w) = P(X \leq w)$$



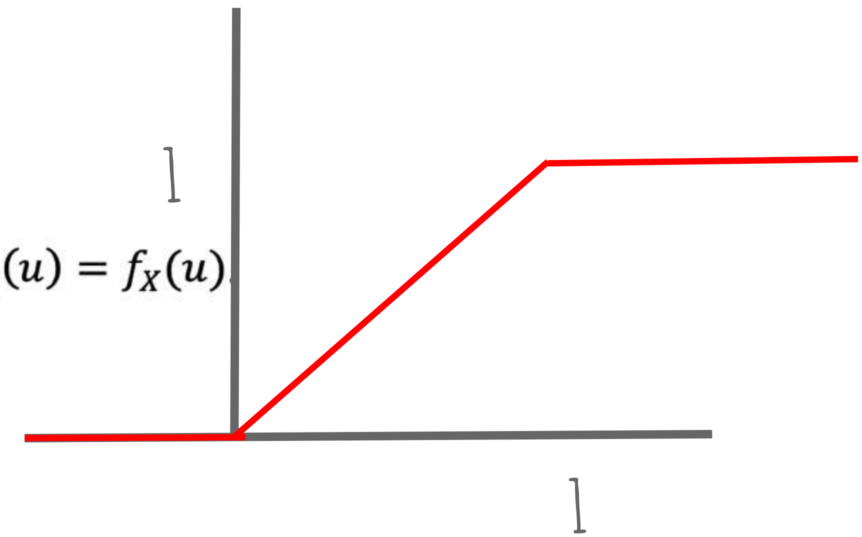
$$F_X(w) = \begin{cases} 0, & w < 0 \\ w, & 0 \leq w \leq 1 \\ 1, & w > 1 \end{cases}$$

CDF Intuition

$$F_X(w) = P(X \leq w)$$

$F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f_X(w)dw$ for all $t \in \mathbb{R}$.

Hence, by the Fundamental Theorem of Calculus, $\frac{d}{du} F_X(u) = f_X(u)$.



$$F_X(w) = \begin{cases} 0, & w < 0 \\ w, & 0 \leq w \leq 1 \\ 1, & w > 1 \end{cases}$$

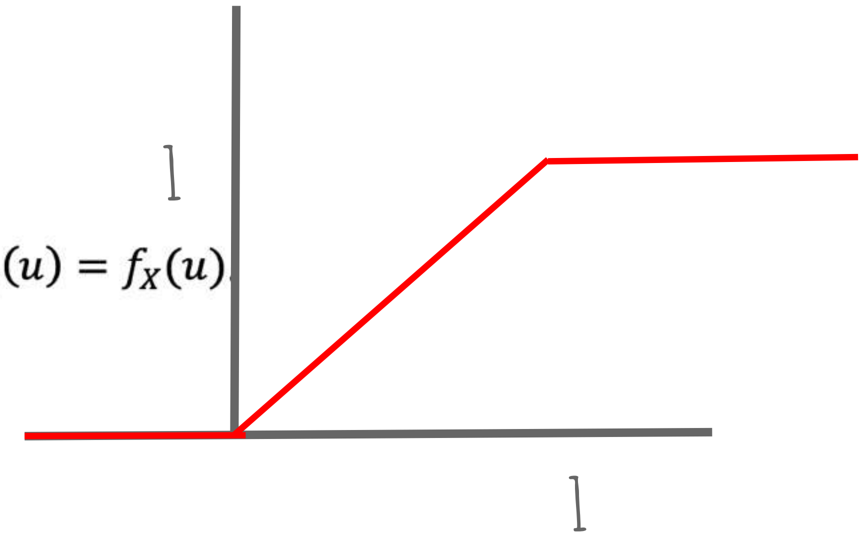
CDF Intuition

$$F_X(w) = P(X \leq w)$$

$F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f_X(w)dw$ for all $t \in \mathbb{R}$.

Hence, by the Fundamental Theorem of Calculus, $\frac{d}{du}F_X(u) = f_X(u)$.

$P(a \leq X \leq b) = F_X(b) - F_X(a)$.



$$F_X(w) = \begin{cases} 0, & w < 0 \\ w, & 0 \leq w \leq 1 \\ 1, & w > 1 \end{cases}$$

CDF Intuition

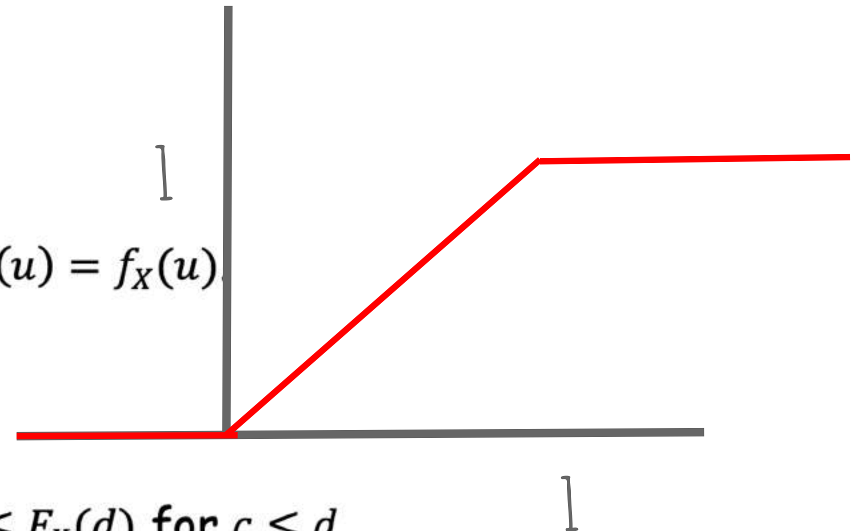
$$F_X(w) = P(X \le w)$$

$F_X(t) = P(X \le t) = \int_{-\infty}^{t} f_X(w)dw$ for all $t \in \mathbb{R}$.

Hence, by the Fundamental Theorem of Calculus, $\frac{d}{du}F_X(u) = f_X(u)$.

$P(a \le X \le b) = F_X(b) - F_X(a)$.

F_X is monotone increasing, since $f_X \ge 0$. That is, $F_X(c) \le F_X(d)$ for $c \le d$.



$$F_X(w) = \begin{cases} 0, & w < 0 \\ w, & 0 \le w \le 1 \\ 1, & w > 1 \end{cases}$$

CDF Intuition

$F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f_X(w)dw$ for all $t \in \mathbb{R}$.

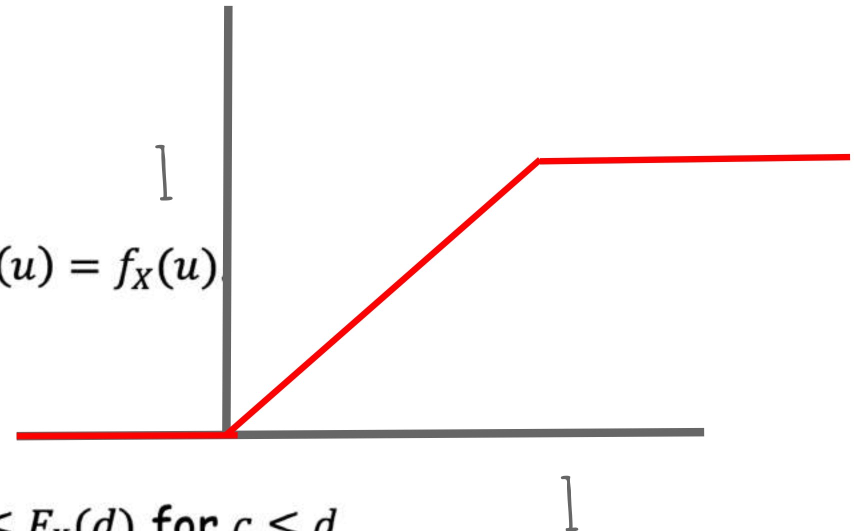
Hence, by the Fundamental Theorem of Calculus, $\frac{d}{du}F_X(u) = f_X(u)$.

$P(a \leq X \leq b) = F_X(b) - F_X(a)$.

F_X is monotone increasing, since $f_X \geq 0$. That is, $F_X(c) \leq F_X(d)$ for $c \leq d$.

$\lim_{v \to -\infty} F_X(v) = P(X \leq -\infty) = 0$.

$$F_X(w) = P(X \leq w)$$



$$F_X(w) = \begin{cases} 0, & w < 0 \\ w, & 0 \leq w \leq 1 \\ 1, & w > 1 \end{cases}$$

CDF Intuition

$$F_X(w) = P(X \leq w)$$

$F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f_X(w)dw$ for all $t \in \mathbb{R}$.

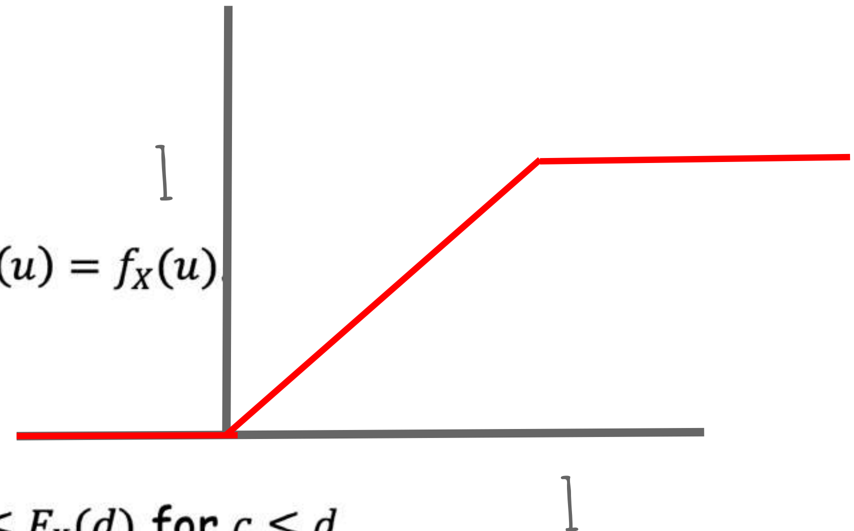
Hence, by the Fundamental Theorem of Calculus, $\frac{d}{du}F_X(u) = f_X(u)$.

$P(a \leq X \leq b) = F_X(b) - F_X(a)$.

F_X is monotone increasing, since $f_X \geq 0$. That is, $F_X(c) \leq F_X(d)$ for $c \leq d$.

$\lim_{v \to -\infty} F_X(v) = P(X \leq -\infty) = 0$.

$\lim_{v \to +\infty} F_X(v) = P(X \leq +\infty) = 1$.



$$F_X(w) = \begin{cases} 0, & w < 0 \\ w, & 0 \leq w \leq 1 \\ 1, & w > 1 \end{cases}$$

Cumulative Distribution Functions (CDFs)

Cumulative Distribution Function (CDF): Let X be a continuous rv (one whose range is typically an interval or union of intervals). The cumulative distribution function (CDF) of X is the function $F_X: \mathbb{R} \to \mathbb{R}$ such that

- $F_X(t) = P(X \leq t) = \int_{-\infty}^{t} f_X(w)dw$ for all $t \in \mathbb{R}$.
- Hence, by the Fundamental Theorem of Calculus, $\frac{d}{du}F_X(u) = f_X(u)$.
- $P(a \leq X \leq b) = F_X(b) - F_X(a)$.
- F_X is monotone increasing, since $f_X \geq 0$. That is, $F_X(c) \leq F_X(d)$ for $c \leq d$.
- $\lim_{v \to -\infty} F_X(v) = P(X \leq -\infty) = 0$.
- $\lim_{v \to +\infty} F_X(v) = P(X \leq +\infty) = 1$.

From Discrete to Continuous

	Discrete	Continuous
PMF/PDF	$p_X(x) = P(X = x)$	$f_X(x) \neq P(X = x) = 0$
CDF	$F_X(x) = \sum_{t \leq x} p_X(t)$	$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt$
Normalization	$\sum_x p_X(x) = 1$	$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$
Expectation	$\mathbb{E}[g(X)] = \sum_x g(x) p_X(x)$	$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\, dx$

Probability

4.2 Zoo of Continuous RVs

Agenda

- The (Continuous) Uniform RV
- The Exponential RV
- Memorylessness

The (Continuous) Uniform RV

<u>**Uniform (Continuous) RV:**</u> $X \sim Unif(a, b)$ where $a < b$ are real numbers, if and only if X has the following pdf:

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

X is equally likely to take on any value in $[a, b]$.

The Uniform (Continuous) RV

Uniform (Continuous) RV: $X \sim Unif(a, b)$ where $a < b$ are real numbers, if and only if X has the following pdf:

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

X is equally likely to take on any value in $[a, b]$.

$$E[X] = \frac{a+b}{2} \qquad Var(X) = \frac{(b-a)^2}{12}$$

The cdf is

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

The Exponential PDF/CDF

Recall the Poisson Process with parameter $\lambda > 0$ has events happening at average rate of λ per unit of time forever. The exponential RV measures the time until the first occurrence of an event, so is a continuous RV with range $[0, \infty)$ (unlike the Poisson RV, which counts the number of occurrences in a unit of time, with range $\{0,1,2,\ldots\}$.)

The Exponential PDF/CDF

Recall the Poisson Process with parameter $\lambda > 0$ has events happening at average rate of λ per unit of time forever. The exponential RV measures the time until the first occurrence of an event, so is a continuous RV with range $[0, \infty)$ (unlike the Poisson RV, which counts the number of occurrences in a unit of time, with range $\{0,1,2,\ldots\}$.)

Let $Y \sim Exp(\lambda)$ be the time until the first event. We'll compute $F_Y(t)$ and $f_Y(t)$.
Let $X(t) \sim Poi(\lambda t)$ be the # of events in the first t units of time, for $t \geq 0$.

The Exponential PDF/CDF

Recall the Poisson Process with parameter $\lambda > 0$ has events happening at average rate of λ per unit of time forever. The exponential RV measures the time until the first occurrence of an event, so is a continuous RV with range $[0, \infty)$ (unlike the Poisson RV, which counts the number of occurrences in a unit of time, with range $\{0,1,2, \dots\}$.)

Let $Y \sim Exp(\lambda)$ be the time until the first event. We'll compute $F_Y(t)$ and $f_Y(t)$.
Let $X(t) \sim Poi(\lambda t)$ be the # of events in the first t units of time, for $t \geq 0$.

$$P(Y > t) = P(\text{no events in first } t \text{ units}) = P(X(t) = 0) = e^{-\lambda t} \frac{(\lambda t)^0}{0!} = e^{-\lambda t}$$

$$F_Y(t) = P(Y \leq t) = 1 - P(Y > t) = 1 - e^{-\lambda t}$$

$$f_Y(t) = \frac{d}{dt} F_Y(t) = \lambda e^{-\lambda t}$$

The Exponential RV Properties

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

The Exponential RV Properties

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{0}^{\infty} x \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda}$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_{0}^{\infty} x^2 \cdot \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2}$$

$$Var(X) = E[X^2] - E[X]^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

The Exponential RV

<u>Exponential RV:</u> $X \sim Exp(\lambda)$, if and only if X has the following pdf:

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

X is the waiting time until the first occurrence of an event in a Poisson process with parameter λ.

The Exponential RV

<u>Exponential RV:</u> $X \sim Exp(\lambda)$, if and only if X has the following pdf:

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

X is the waiting time until the first occurrence of an event in a Poisson process with parameter λ.

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

The cdf is

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

Random Picture

Memorylessness (Intuition)

A random variable X is **memoryless** if for all $s, t \geq 0$,

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

Memorylessness (Intuition)

A random variable X is **memoryless** if for all $s, t \geq 0$,

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

For example, let $s = 7, t = 2$. So $P(X > 9 \mid X > 7) = P(X > 2)$. That is, given we've waited 7 minutes, the probability we wait at least 2 more, is the same as the probability we wait at least 2 more from the beginning.

Memorylessness (Intuition)

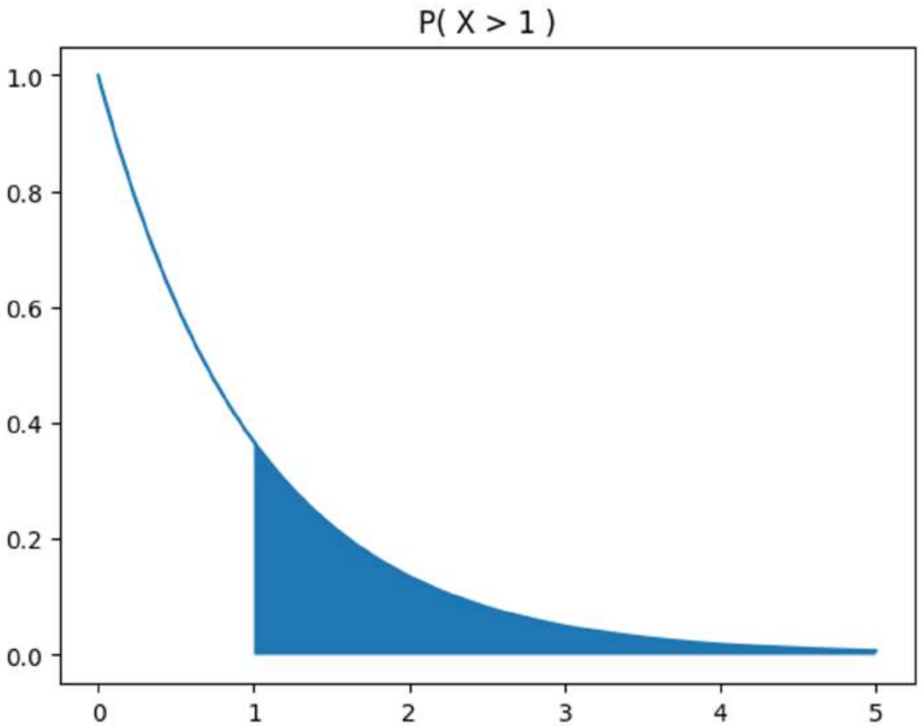
A random variable X is **memoryless** if for all $s, t \geq 0$,

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

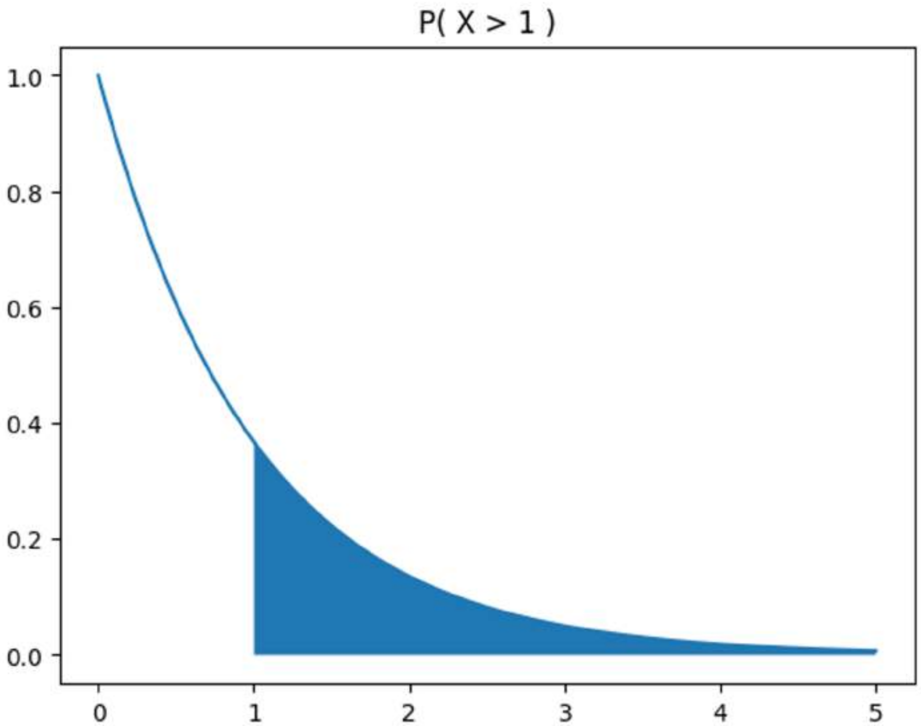
For example, let $s = 7, t = 2$. So $P(X > 9 \mid X > 7) = P(X > 2)$. That is, given we've waited 7 minutes, the probability we wait at least 2 more, is the same as the probability we wait at least 2 more from the beginning.

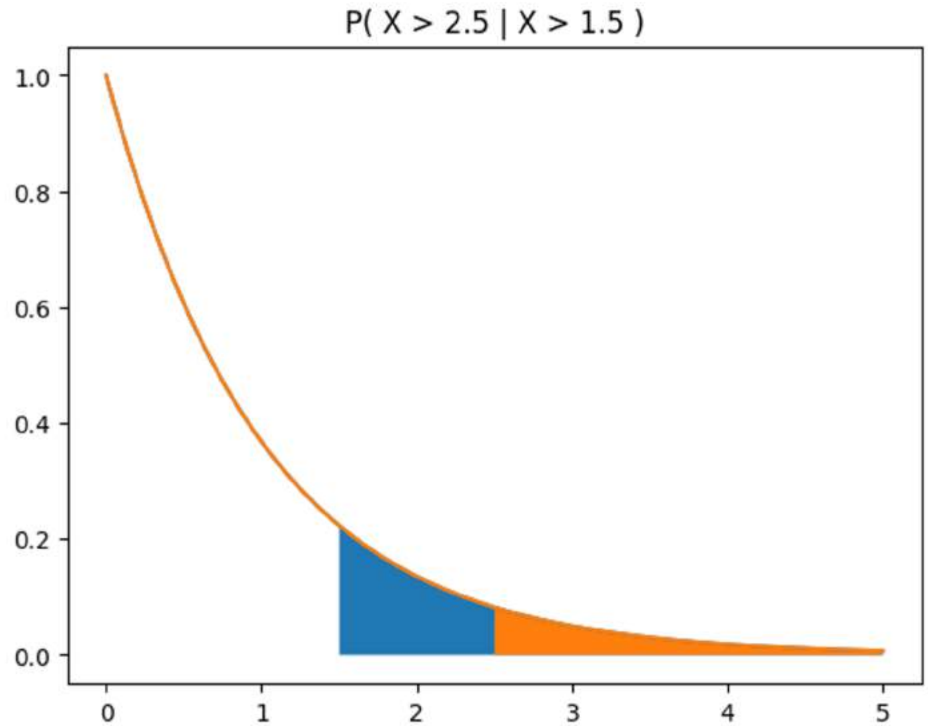
The only memoryless RVs are the **Geometric** (discrete) and **Exponential** (continuous)!

Memorylessness (Intuition)

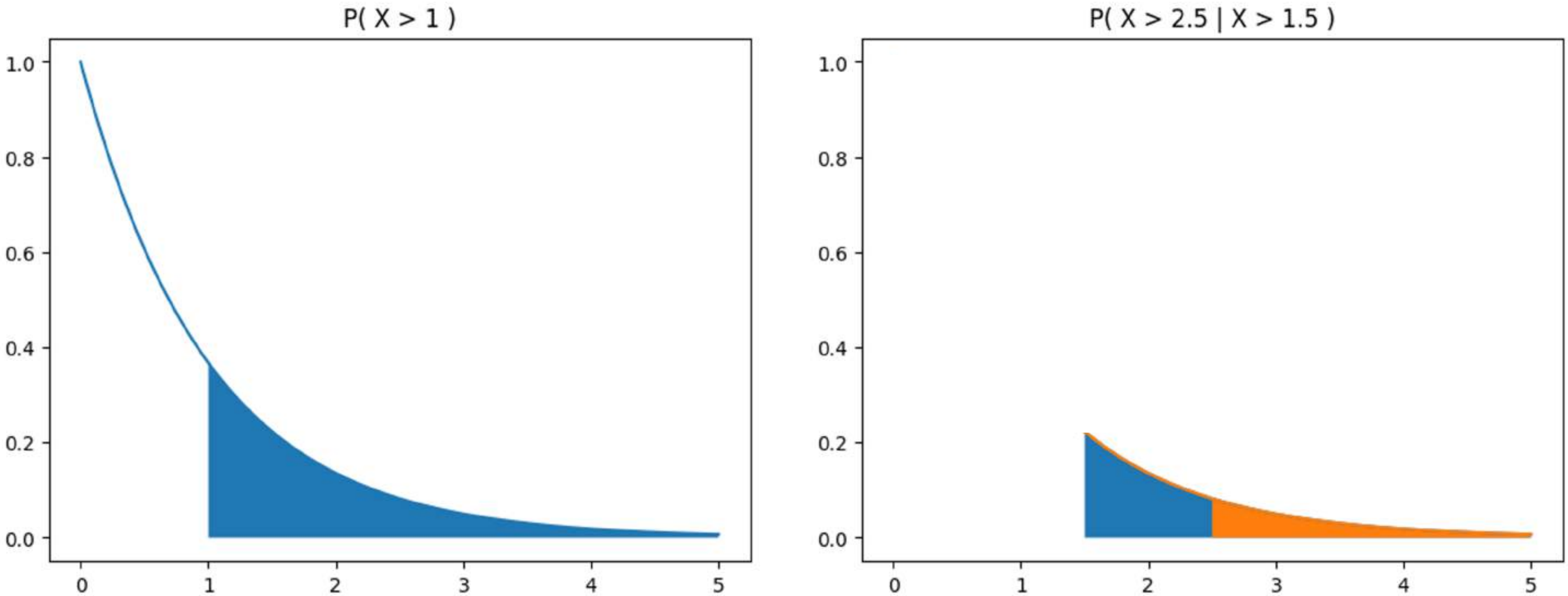


Memorylessness (Intuition)





Memorylessness (Intuition)



Memorylessness of Exponential (Proof)

If $X \sim Exp(\lambda)$ and $x \geq 0$, then recall

$$P(X > x) = 1 - F_X(x) = 1 - (1 - e^{-\lambda x}) = e^{-\lambda x}$$

Memorylessness of Exponential (Proof)

If $X \sim Exp(\lambda)$ and $x \geq 0$, then recall

$$P(X > x) = 1 - F_X(x) = 1 - (1 - e^{-\lambda x}) = e^{-\lambda x}$$

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

Memorylessness of Exponential (Proof)

If $X \sim Exp(\lambda)$ and $x \geq 0$, then recall

$$P(X > x) = 1 - F_X(x) = 1 - (1 - e^{-\lambda x}) = e^{-\lambda x}$$

$$\begin{aligned} P(X > s + t \mid X > s) &= \frac{P(X > s \mid X > s + t)P(X > s + t)}{P(X > s)} \\ &= \frac{P(X > s + t)}{P(X > s)} \\ &= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} \\ &= e^{-\lambda t} \\ &= P(X > t) \end{aligned}$$

The Gamma RV

<u>Gamma RV:</u> $X \sim Gamma(r, \lambda)$ if and only if X has the following pdf:

$$f_X(x) = \begin{cases} \frac{\lambda^r}{(r-1)!} x^{r-1} e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

X is the sum of r independent $Exp(\lambda)$ random variables.