ENGG 2760A / ESTR 2018: Probability for Engineers

7. Continuous Random Variables

Prof. Hong Xu

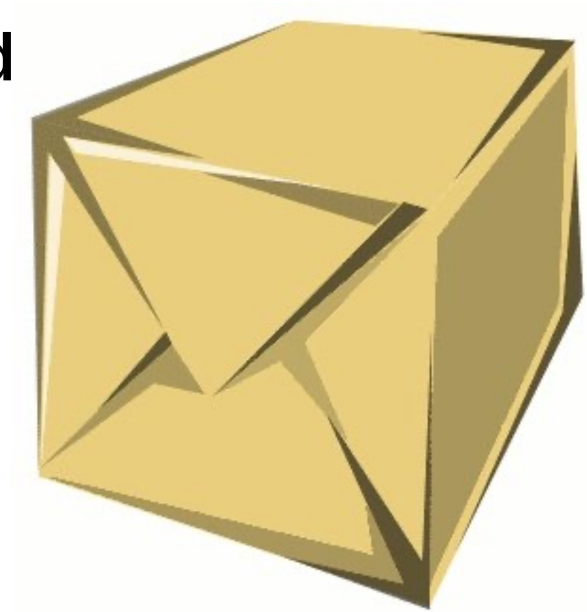
Credit to Prof. Andrej Bogdanov

Delivery time

A package is to be delivered between noon and 1pm.

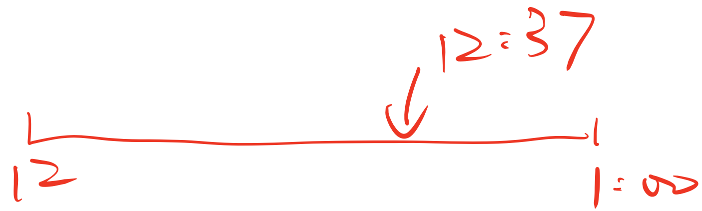
You go out between 12:30 and 12:45.

What is the probability you missed the delivery?

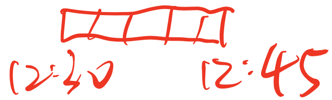


Delivery time

1. Sample space:



2. Event of interest:



3. Probabilities?

$$P(12{:}a \le T < 12{:}b) = \frac{b-a}{60}$$

$$P(E) = \frac{45-30}{60} = \frac{1}{4}$$

Uncountable sample spaces

In Lecture 1 we said:

> "*The probability of an event is the sum of the probabilities of its elements*"

...but all elements have probability zero!

To specify and calculate probabilites, we have to work with the axioms of probability

Delivery time

From 12:08 - 12:12 and 12:54 - 12:57 the doorbell wasn’t working.

Event of interest: $\overset{A}{\boxed{08 \quad 12}} \cup \overset{B}{\boxed{54 \quad 57}}$

Probability: $P(A \cup B) = P(A) + P(B) = \frac{4}{60} + \frac{3}{60} = \frac{7}{60}$

The uniform random variable

Sample space $\Omega = [0, 60)$

Events of interest: intervals $[x, y) \subseteq [0, 60)$

their intersections, unions, etc.

Probabilities: $P([x, y)) = (y - x)/60$

Random variable: $X(\omega) = \omega$

Cumulative distribution function

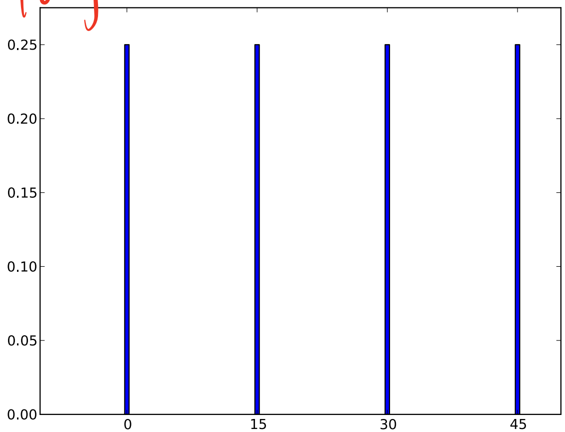
The probability mass function doesn't make much sense because $P(X = x) = 0$ for all x.

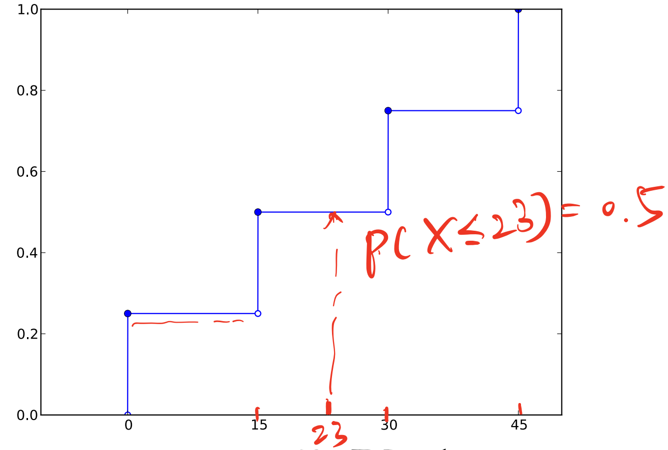
Instead, we can describe X by its cumulative distribution function (CDF) F:

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

Cumulative distribution functions

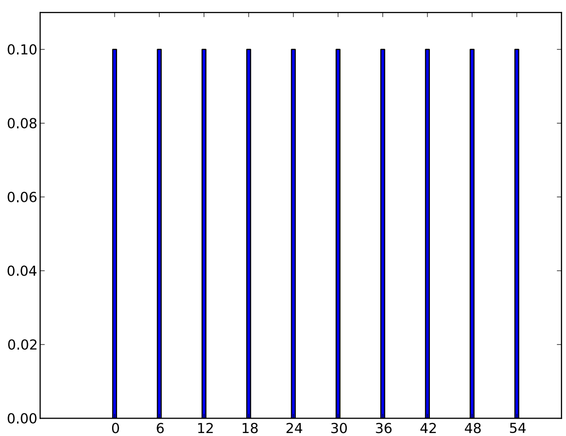
pmf

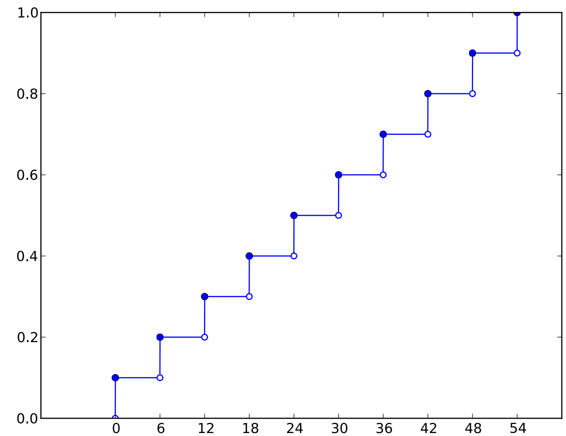




$$f_X(x) = P(X = x)$$

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

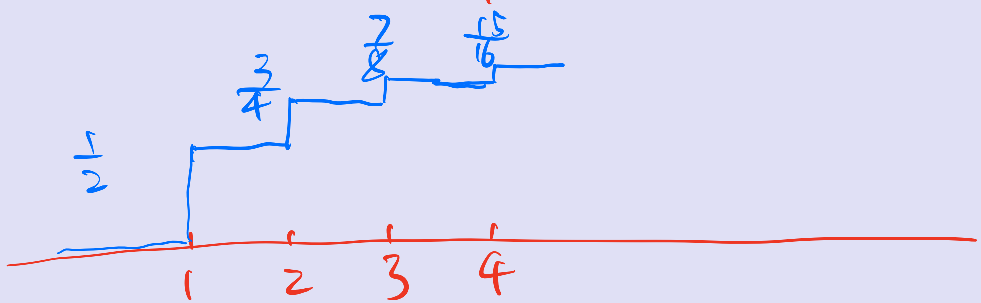




What is the $\text{Geometric}(1/2)$ CDF?

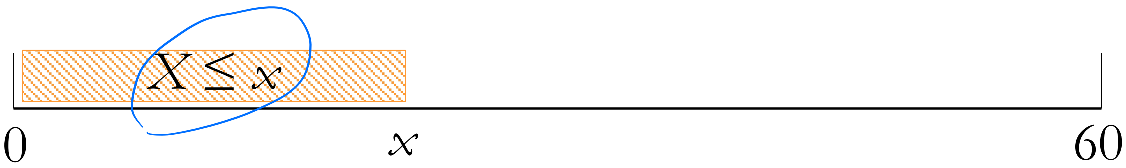
PMF: $P(N=n) = \frac{1}{2^n}$, $n = 1, 2, 3, \cdots$.

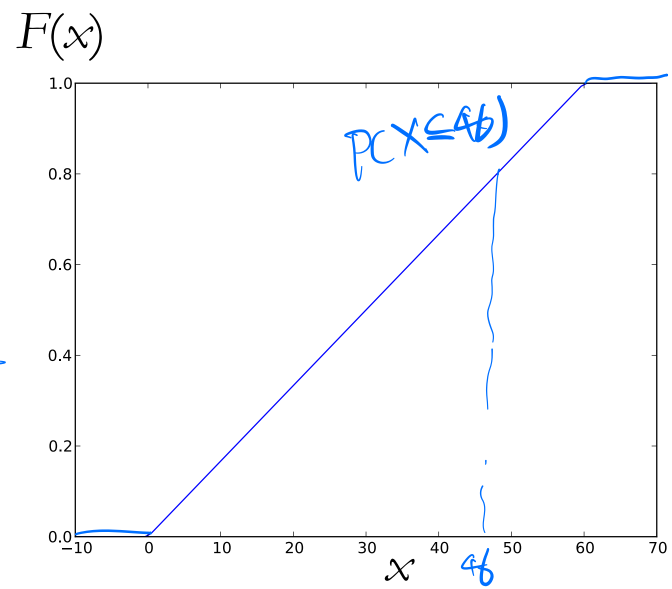
CDF: $P(N \le n) = \sum_{1}^{n} \frac{1}{2^n} = 1 - \frac{1}{2^n}$, $n = 1, 2, 3, \cdots$.



Uniform random variable

If X is uniform over $[0, 60)$ then

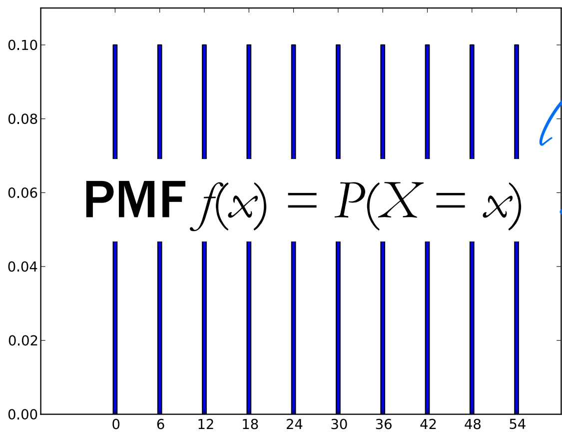


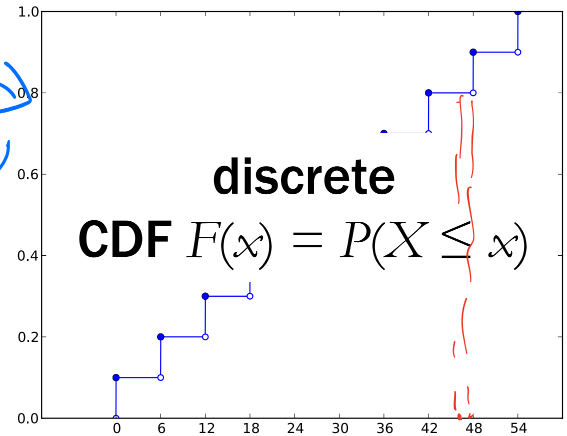


$$P(X \leq x) = \begin{cases} 0 & \textbf{for } x < 0 \\ \boxed{x/60} & \textbf{for } x \in [0, 60) \\ 1 & \textbf{for } x > 60 \end{cases}$$

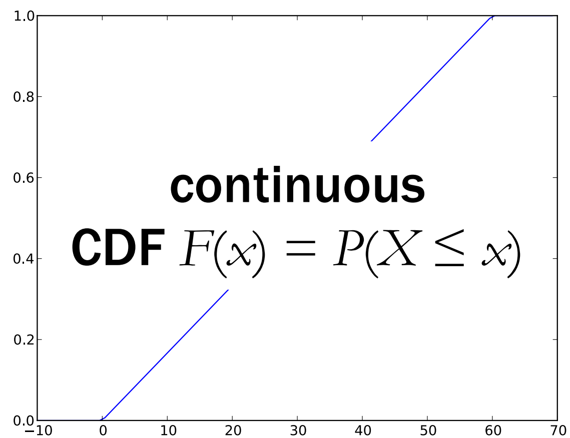
Cumulative distribution functions

$F(48) - F(48-\delta)$
$= 0.9 - 0.8$
$= 0.1$
$= f(48)$





?



Discrete random variables:

PMF $f(x) = P(X = x)$ $\underset{\text{SUBTRACT}}{\overset{\text{ADD}}{\rightleftarrows}}$ **CDF** $F(x) = P(X \leq x)$

$f(x) = F(x) - F(x - \delta)$ $\quad\quad$ $F(a) = \sum_{x \leq a} f(x)$

for small δ

Continuous random variables:

The probability density function (PDF) of a random variable with CDF $F(x)$ is

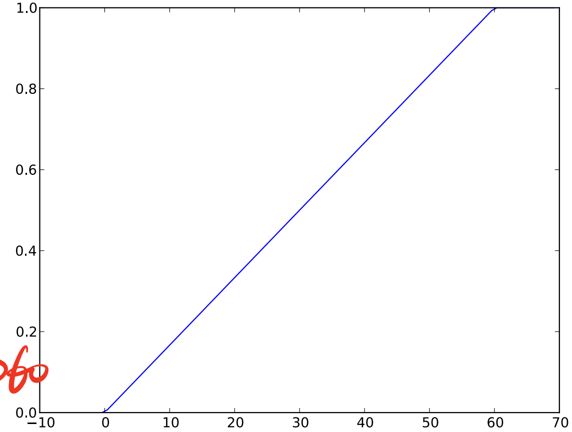
$$f(x) = \lim_{\delta \to 0} \frac{F(x) - F(x - \delta)}{\delta} = \frac{dF(x)}{dx}$$

derivative

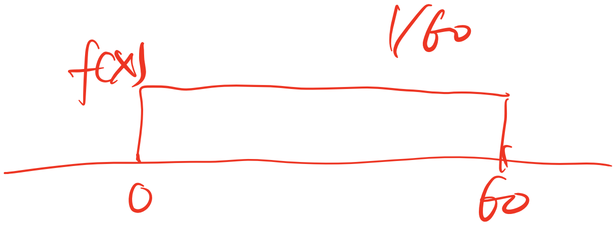
Δ

Uniform random variable

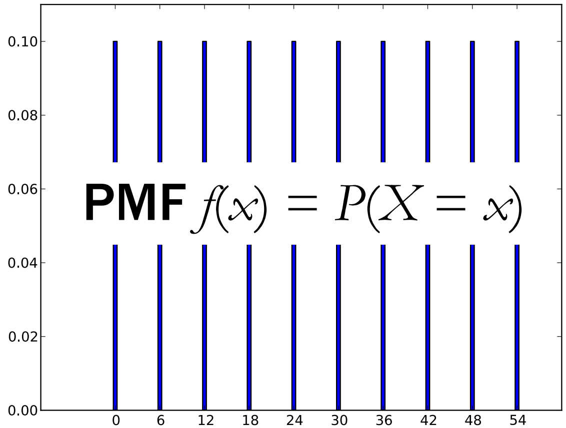
$$F(x) = \begin{cases} 0 & \textbf{if } x < 0 \\ x/60 & \textbf{if } x \in [0, 60) \\ 1 & \textbf{if } x \geq 60 \end{cases}$$

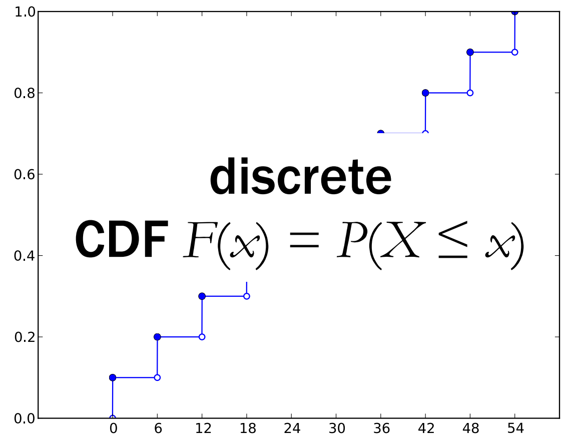


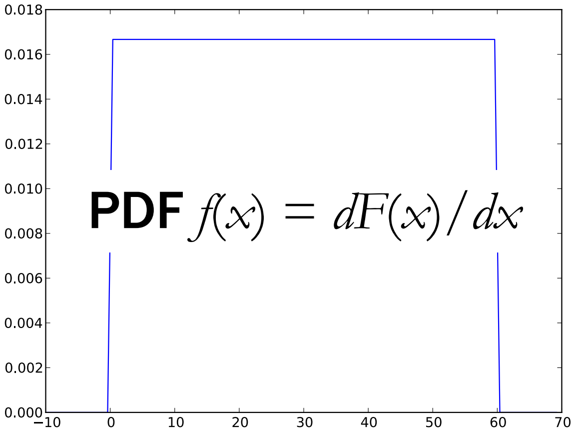
PDF: $f(x) = \dfrac{dF(x)}{dx} = \begin{cases} 0, & x<0 \text{ OR } x>60 \\ \frac{1}{60}, & x \in [0, 60]. \end{cases}$

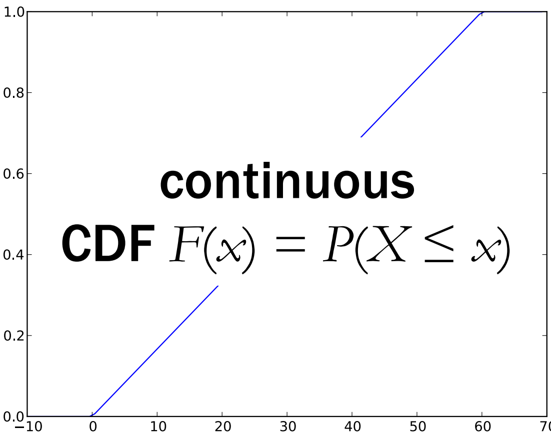


Probability density functions







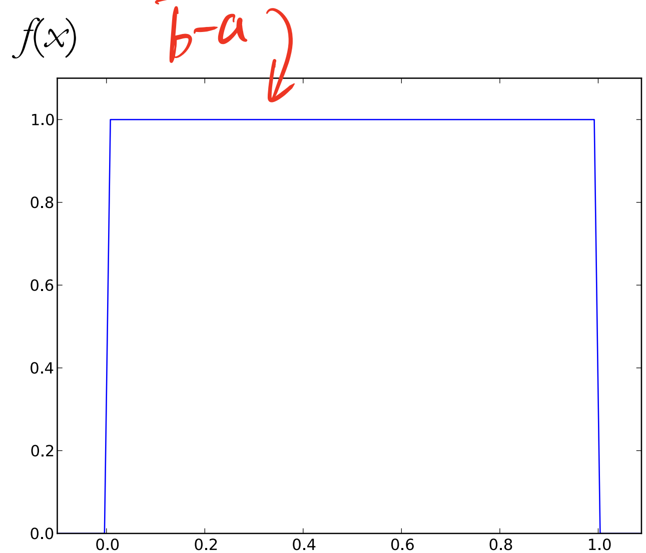


Uniform random variable

The $\mathrm{Uniform}(0, 1)$ **PDF is**

$$f(x) = \begin{cases} 1 & \textbf{if } x \in (0, 1) \\ 0 & \textbf{if } x < 0 \textbf{ or } x > 1 \end{cases}$$

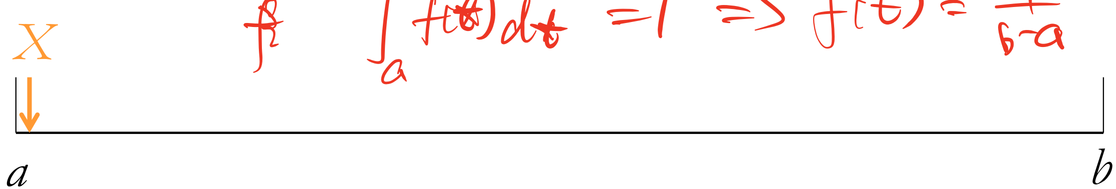
$\frac{1}{b-a}$



The $\mathrm{Uniform}(a, b)$ **PDF is**

$$f(x) = \begin{cases} 1/(b - a) & \textbf{if } x \in (a, b) \\ 0 & \textbf{if } x < a \textbf{ or } x > b \end{cases}$$

$\int_a^b f(t)\,dt = 1 \Rightarrow f(t) = \frac{1}{b-a}$



Calculating the CDF

Discrete random variables:

PMF $f(x) = P(X = x)$

CDF $F(x) = P(X \le x) = \sum_{t \le x} f(t)$

Continuous random variables:

PDF $f(x) = dF(x)/dx$

CDF $F(x) = P(X \le x) = \int_{t \le x} f(t)\,dt$

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

A package is to arrive between 12 and 1

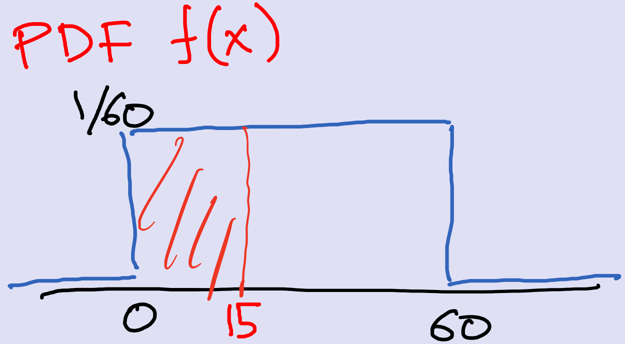
What is the probability it arrived by 12.15?

$X \sim \text{Uniform}$

$$P(X \le 15) = \int_{-\infty}^{15} f(t)\,dt$$

$$= \frac{t}{60}\Big|_0^{15}$$

$$= \frac{1}{4}.$$



Interpretation of the PDF

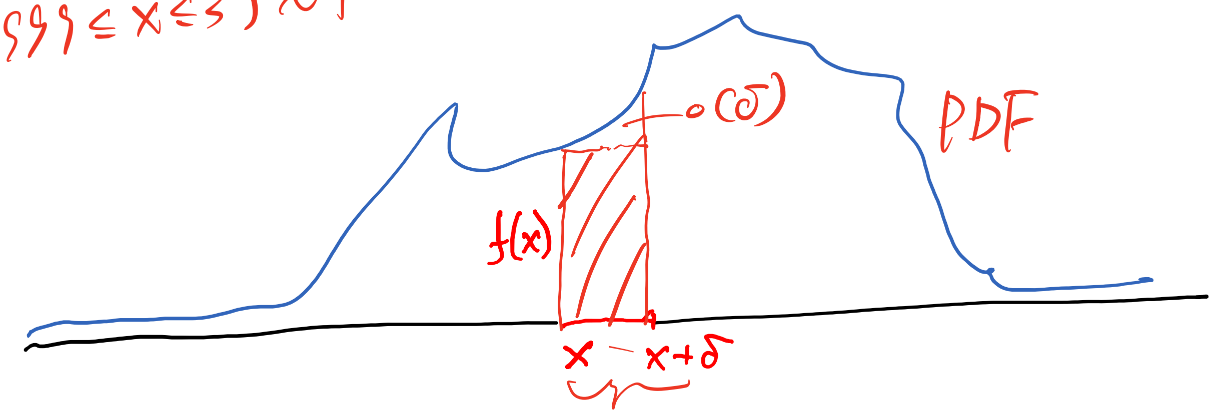
The PDF value $f(x)\ \delta$ approximates the probability that X in an interval of length δ around x

$$P(x-\delta \le X < x) = f(x)\ \delta + o(\delta)$$

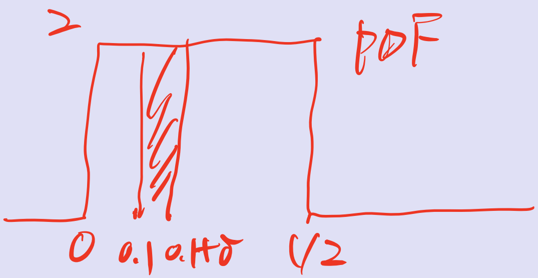
$$P(x \le X < x+\delta) = f(x)\ \delta + o(\delta)$$

$P(X=x)=0$

$P(2.999 \le X \le 3) \approx f(3)\cdot 0.001$



Uniform $(0, \frac{1}{2})$



$\delta = 0.001$

$P(X = 0.1) = 0.$

$P(0.1 \le X \le 0.1 + \delta) \approx 2 \cdot 0.001 = 0.002.$

Alice said she'll show up between 7 and 8, probably around 7.30.

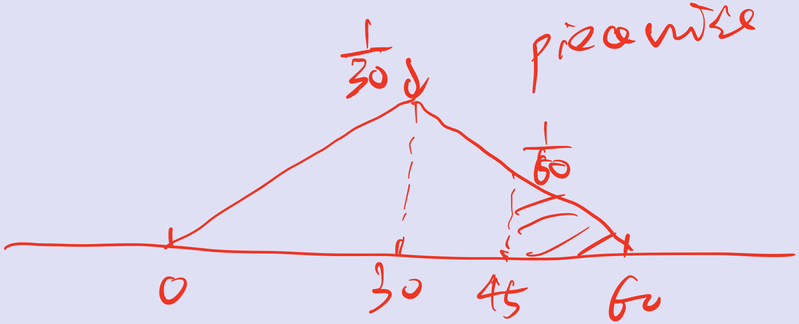
It is now 7.30. What is the probability Bob has to wait past 7.45?

Model:

① A: Alice arrival time, $U(0, 60)$.

② $A \sim$ following PDF:

piecewise linear.



$$P(A \geq 45 \mid A \geq 30)$$

$$= \frac{P(A \geq 45)}{P(A \geq 30)}$$

$$= \frac{\overbrace{\frac{1}{2} \cdot 15 \cdot \frac{1}{60}}^{\frac{1}{8}}}{1/2} = \frac{1}{4}.$$

$P(X \le x)$

$$\underbrace{P(X \le x+\delta)} > \underbrace{P(X \le x)}$$

Expectation and variance

	PMF $f(x)$	PDF $f(x)$
$\mathbf{P}(X \leq a)$ CDF	$\sum_{x \leq a} f(x)$	$\int_{-\infty}^{\infty} f(x)\,dx$
$\mathbf{E}[X]$	$\sum_x x\, f(x)$	$\int_{-\infty}^{\infty} x \cdot f(x)\,dx$
$\mathbf{E}[X^2]$	$\sum_x x^2 f(x)$	$\int_{-\infty}^{\infty} x^2 f(x)\,dx$
$\mathbf{Var}[X]$ $= E[X^2] - E[X]^2$.		

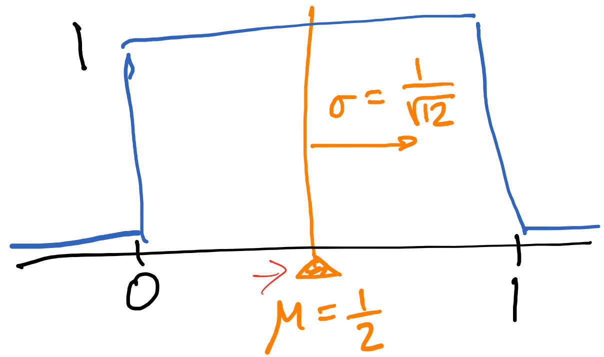
Mean and Variance of Uniform

Uniform(0,1)

$$\mu = E[X] = \int_0^1 x \cdot 1\,dx = \frac{x^2}{2}\Big|_0^1 = \frac{1}{2}$$

$$E[X^2] = \int_0^1 x^2 \cdot 1\,dx = \frac{x^3}{3}\Big|_0^1 = \frac{1}{3}$$

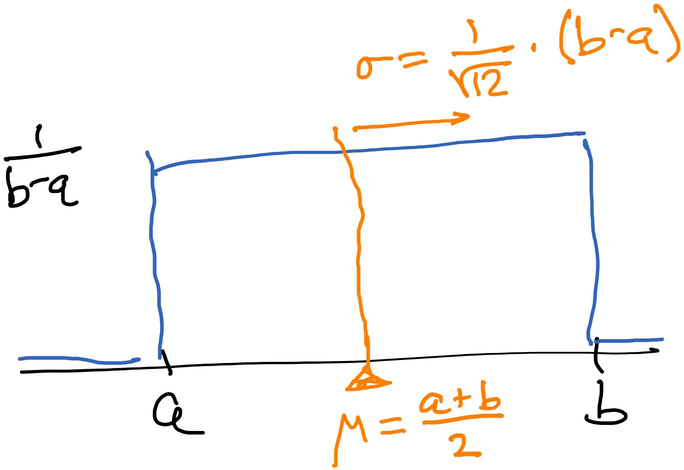
$$\sigma = \sqrt{Var[X]} = \sqrt{1/3 - (1/2)^2} = 1/\sqrt{12}.$$



Uniform(a,b)

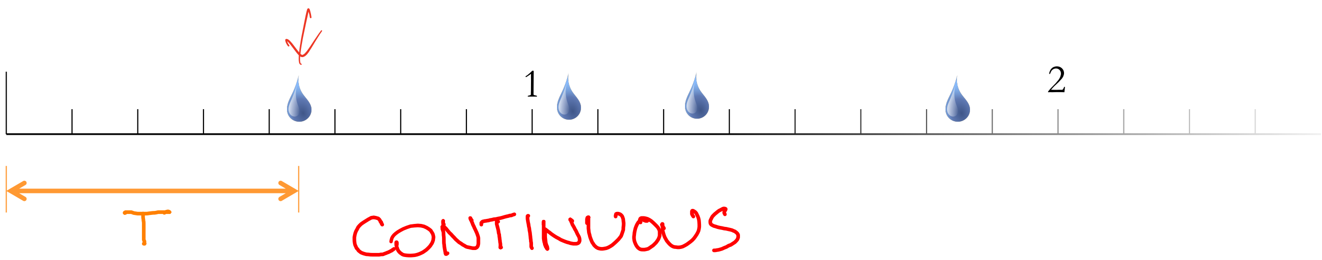
$$\mu = \frac{a+b}{2}$$

$$\sigma = \frac{b-a}{\sqrt{12}}.$$

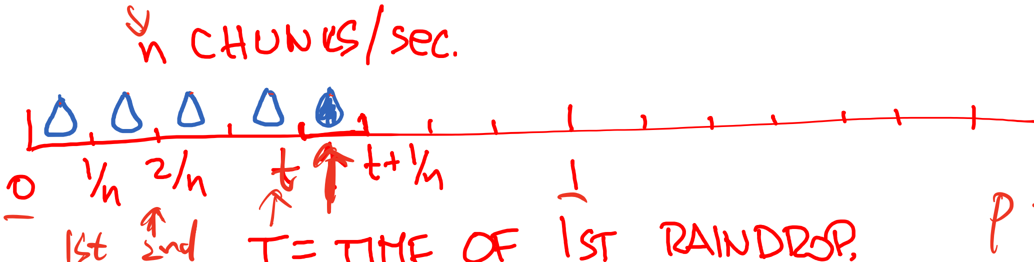


Raindrops again

Rain is falling on your head at a rate of λ drops/sec.



How long do we wait until the next drop?



T = TIME OF 1ST RAINDROP.

$\frac{1}{n}$

$p = \frac{\lambda}{n}$

$t = \frac{1}{n}, \quad tn = \frac{1}{\lambda}$

$$P\left(t \le T < t + \frac{1}{n}\right) = (1-p)^{tn} \cdot p.$$

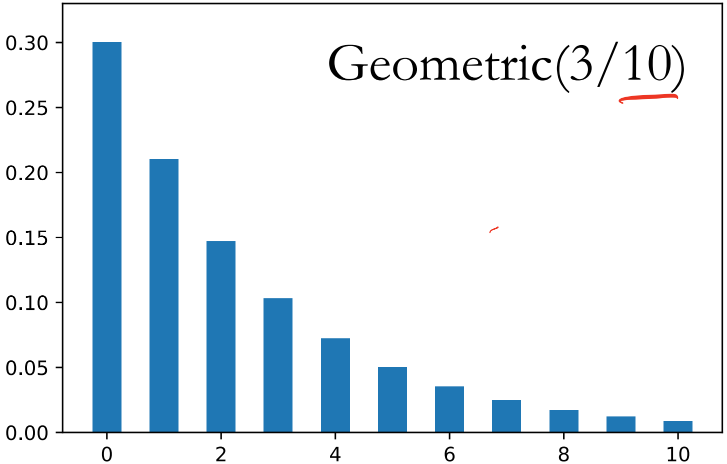
$$= \left(1 - \frac{\lambda}{n}\right)^{tn} \cdot \frac{\lambda}{n}$$

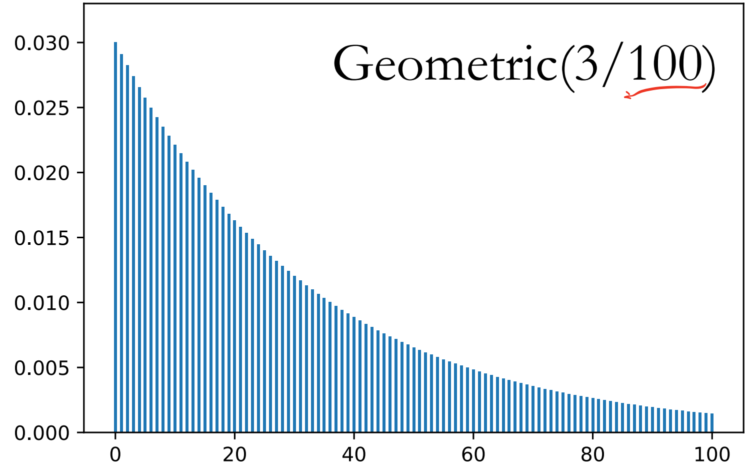
$\delta = \frac{1}{n}$

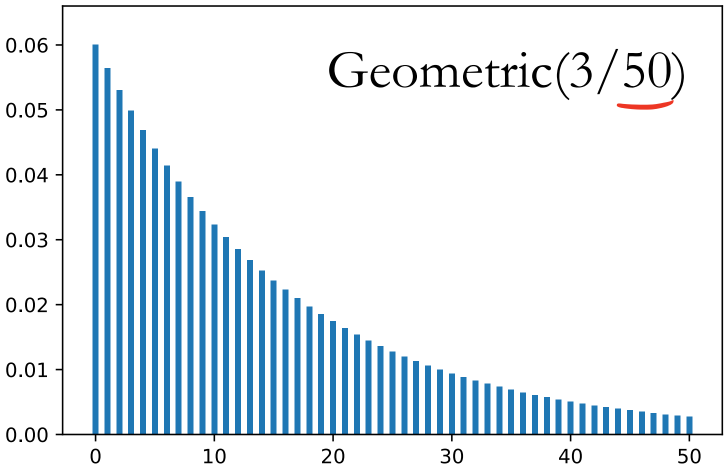
$$P(t \le T < t+\delta) = (1-\lambda\delta)^{t/\delta} \cdot \lambda\delta$$

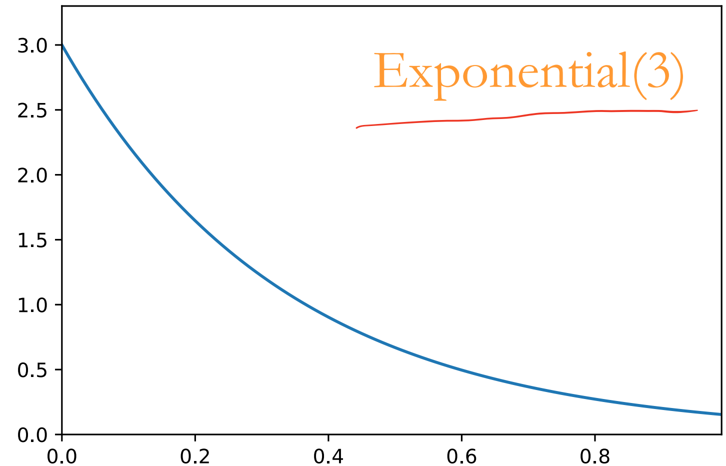
PDF: $$f(t) = \lim_{\delta \to 0} \frac{P(t \le T < t+\delta)}{\delta} = \lim_{\delta \to 0} \lambda (1-\lambda\delta)^{t/\delta}$$

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


Geometric(3/10)
0.30
0.25
0.20
0.15
0.10
0.05
0.00
0
2
4
6
8
10


Geometric(3/100)
0.030
0.025
0.020
0.015
0.010
0.005
0.000
0
20
40
60
80
100

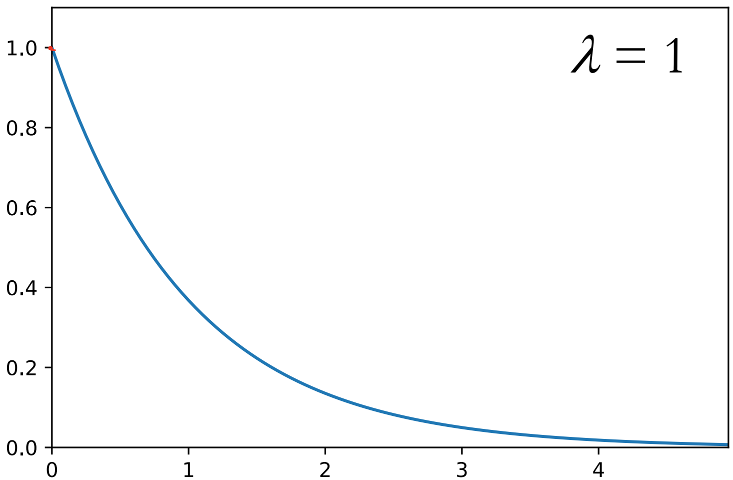

Geometric(3/50)
0.06
0.05
0.04
0.03
0.02
0.01
0.00
0
10
20
30
40
50


Exponential(3)
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0.0
0.2
0.4
0.6
0.8

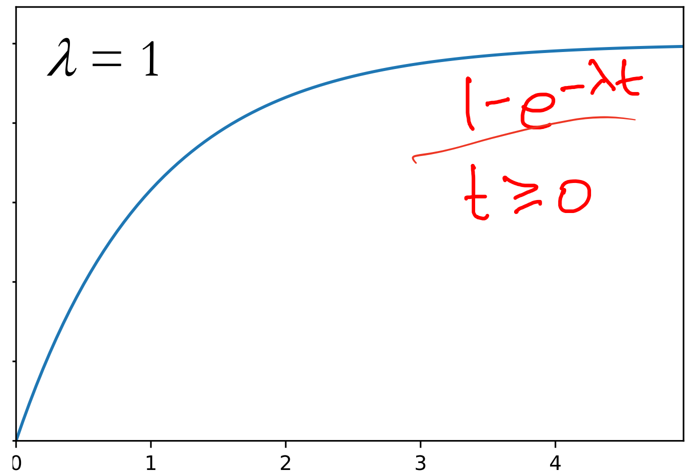
The exponential random variable

The $\text{Exponential}(\lambda)$ **PDF is**

$$f(t) = \begin{cases} \lambda\, e^{-\lambda t} & \textbf{if } \overset{t}{\cancel{x}} \geq 0 \\ 0 & \textbf{if } \underset{t}{\cancel{x}} < 0. \end{cases}$$



PDF $f(t)$



CDF $F(t) = P(T \leq t)$

The exponential random variable

CDF of Exponential(λ)**:** $P(T \leq t) = \int_{-\infty}^{t} f(x)\,dx$

$$= \int_0^t \lambda e^{-\lambda x} \cdot dx$$

$$= -e^{-\lambda x}\Big|_0^t$$

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

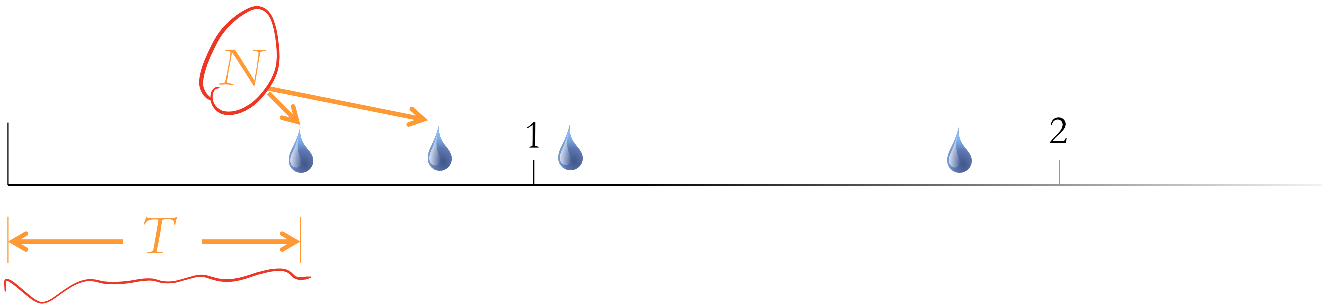
$\mathbf{E}$[Exponential(λ)] $= \lim_{n\to\infty} \frac{1}{n} \cdot \frac{n}{\lambda}$

$$= \boxed{\frac{1}{\lambda}}.$$

Var[Exponential(λ)] $= \lim_{n\to\infty} \frac{1}{n^2} \frac{1-\lambda/n}{(\lambda/n)^2} = \frac{1}{\lambda^2}$. $\sigma = \frac{1}{\lambda}$.

Poisson vs. exponential

Binomial vs. Geometric



	$\text{Poisson}(\lambda)$ N	$\text{Exponential}(\lambda)$ T
description	**number of events within time unit** 0, 1, 2, ...	**time until first event happens** $[0, +\infty)$
expectation	λ	$1/\lambda$
std. deviation	λ	$1/\lambda$

on average

A bus arrives once every 5 minutes. How likely are you to wait 5 to 10 minutes?

$\lambda = \frac{1}{5}$ (bus/min). $E[T] = \frac{1}{\lambda} = 5.$

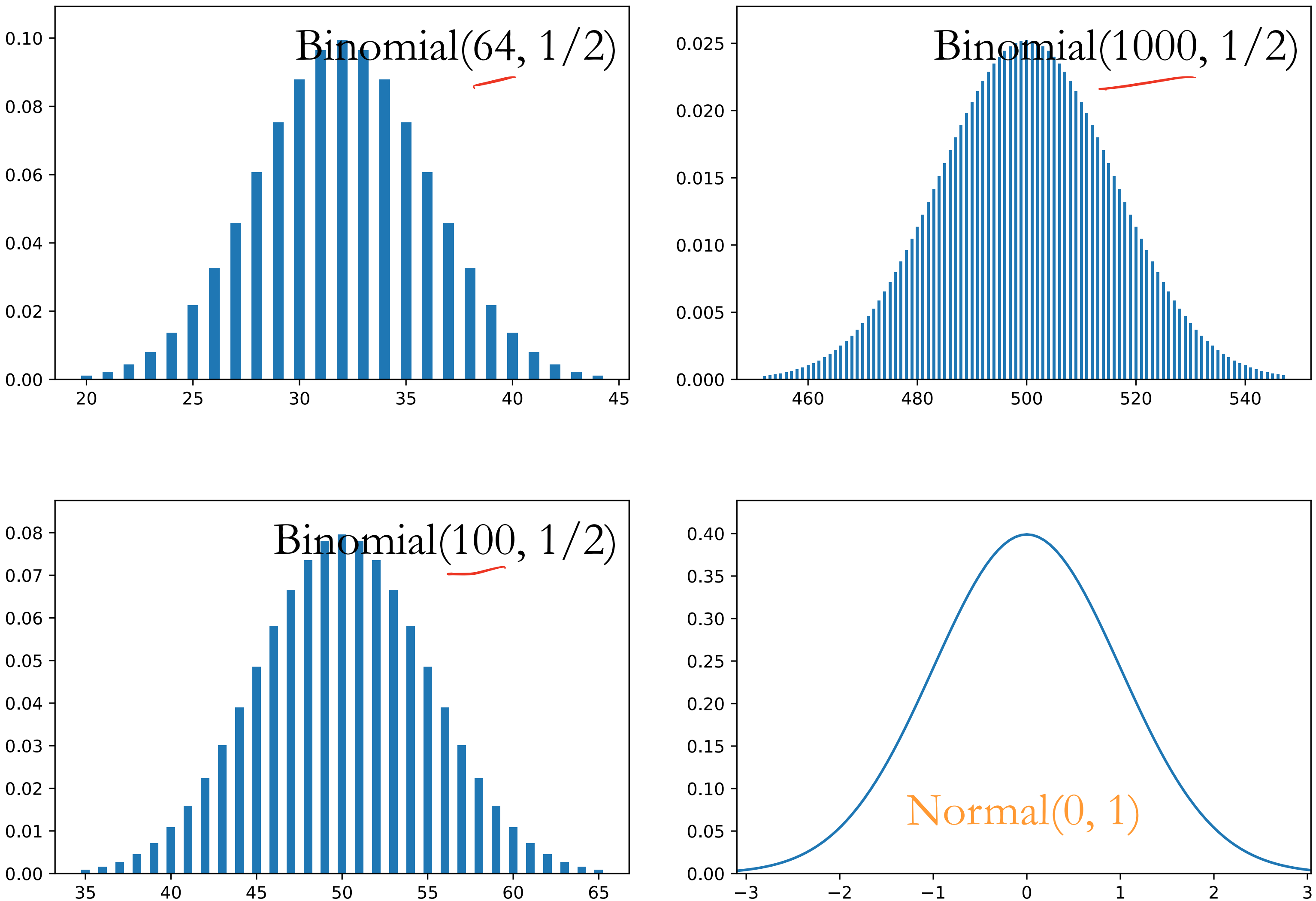
$$P(5 \leq T \leq 10) = P(T \leq 10) - P(T \leq 5)$$

$$\cancel{(5 < T < 10)} = (1 - e^{-\lambda \cdot 10}) - (1 - e^{-\lambda \cdot 5})$$

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

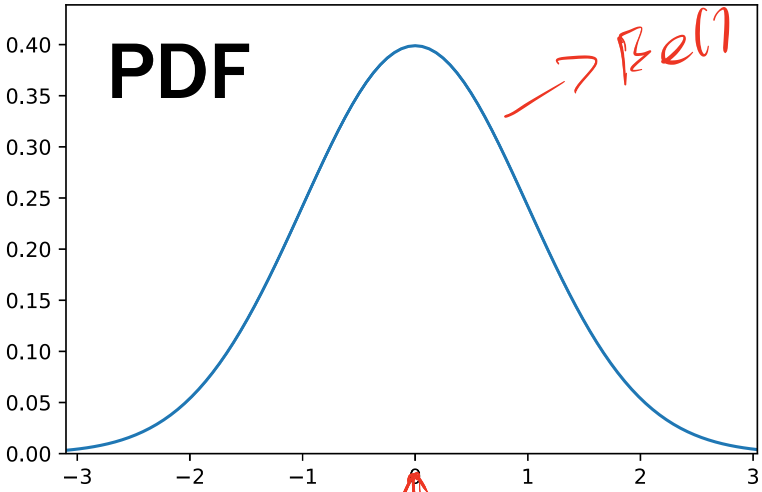
$$P(5 \leq T \leq 10) = \int_5^{10} f(x)\,dx = \int_5^{10} \frac{1}{5} \cdot e^{-x/5}\,dx = \frac{1}{e} - \frac{1}{e^2}$$

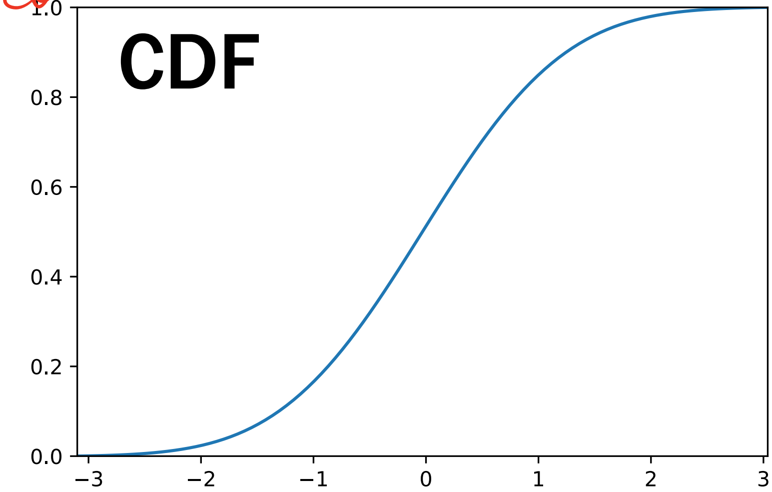
Poisson:


Binomial(64, 1/2)
Binomial(1000, 1/2)
Binomial(100, 1/2)
Normal(0, 1)

The Normal(0, 1) random variable

Gaussian





$$f(x) = (2\pi)^{-1/2} e^{-x^2/2}$$

$$F(x) = (2\pi)^{-1/2} \int_{t \le x} e^{-t^2/2} \, dt$$

no closed form

$$\mathbf{E}[\text{Normal}(0, 1)] = 0$$

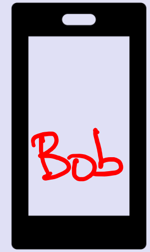
$$\mathbf{Var}[\text{Normal}(0, 1)] = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} x^2 e^{-x^2/2} dx - 0 = 1.$$

$Y = X + N$

Alice

$$-1 \longrightarrow -1 + N$$

$$1 \longrightarrow 1 + N$$



$$- \longmapsto -1$$

$$+ \longmapsto +1$$

$$-1 \longrightarrow \begin{array}{l} -0.73 \\ -1.11 \\ +0.15 \end{array}$$

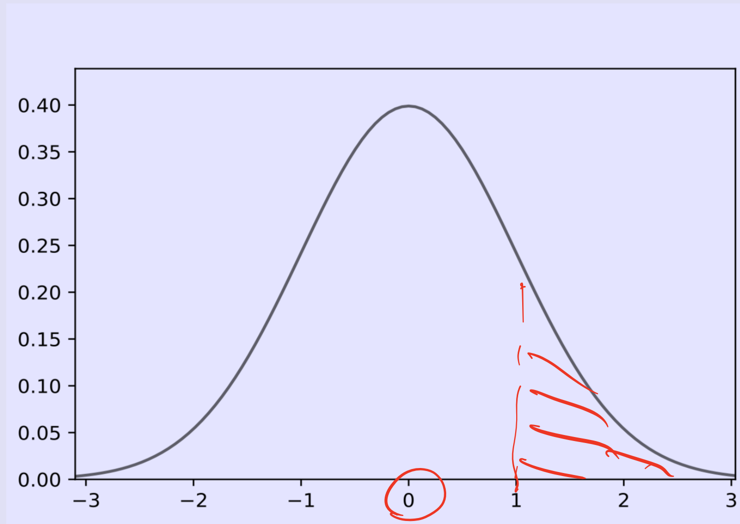
$$+0.15 \longmapsto \text{ERROR}$$

$$+1 \longrightarrow \begin{array}{l} +0.89 \\ -0.031 \end{array}$$

$$-0.031 \longrightarrow \text{ERROR}$$

$$P(\text{ERROR}) = P(N > 1)$$

$$= \int_1^\infty f_N(t)\, dt$$



https://stattrek.com/online-calculator/normal.aspx

The Normal(μ, σ) random variable

$$f(x) = (2\pi\sigma^2)^{-1/2}\, e^{-(x-\mu)^2/2\sigma^2}$$

$E[X] = \mu \qquad Var[X] = \sigma^2$

$$\boxed{\text{Normal}(\mu, \sigma) = \mu + \sigma \cdot \text{Normal}(0, 1)}$$