1179: Probability
Lecture 14 — Special Continuous Random Variables

Ping-Chun Hsieh (謝秉均)

October 29, 2021

Announcements

- TA Hour: 11/5 (Fri.), 6:30pm-8pm @ EC513

Quick Review

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

- f_X $\quad$ $F_X(t)$
- PDF ⇔ CDF?

$f_X(x) = F_X'(x)$ (continuity of $f_X(x)$)

- Continuous uniform r.v., i.e. Unif(a, b)?

$X \sim Unif(a,b)$

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

- Inverse transform sampling (ITS)?

Given $F(t)$, generate a random variable X with CDF $F(t)$

ITS: step1: $U \sim Unif(0,1)$

step2: $X = F^{-1}(U)$

This Lecture

1. Special Continuous Random Variables

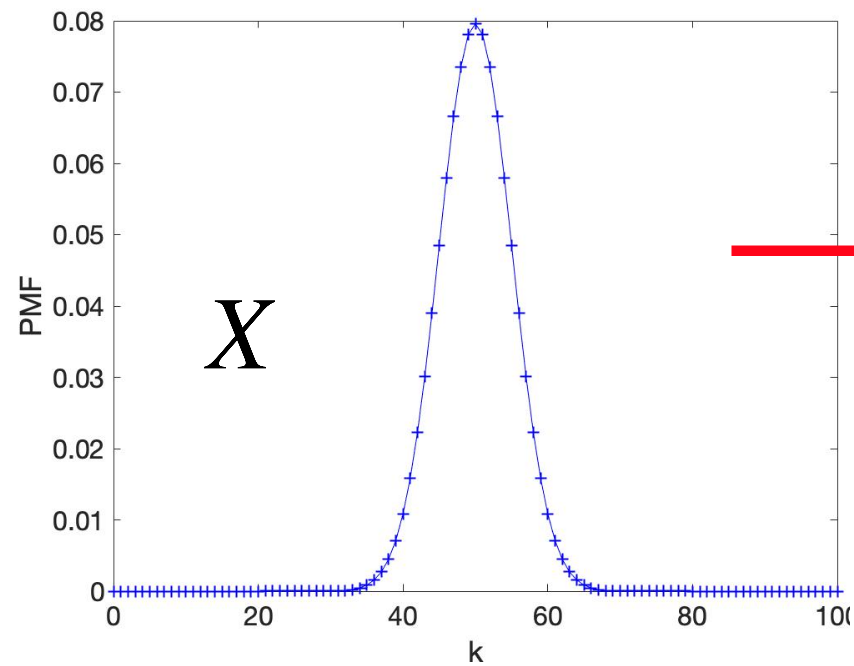
- Reading material: Chapter 7.2~7.3

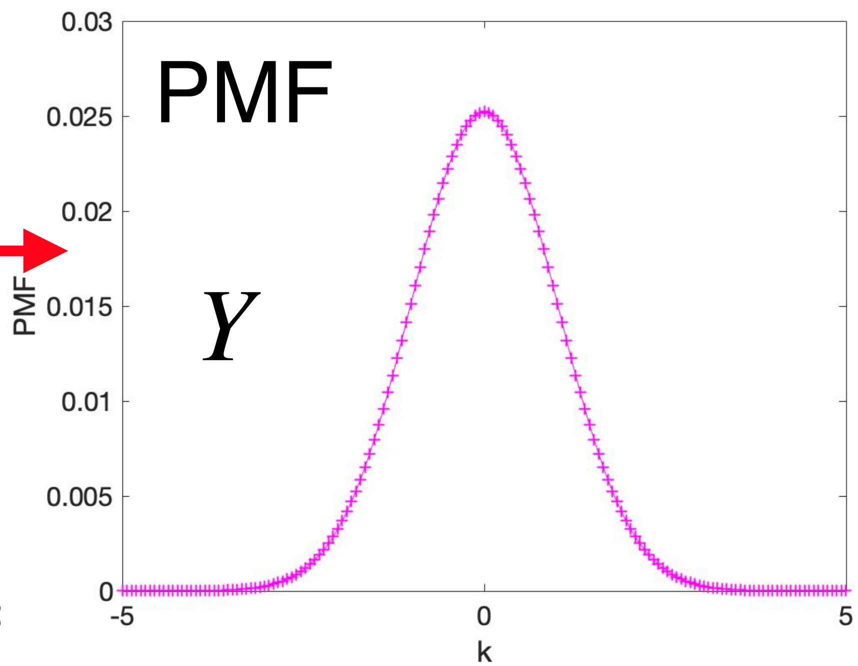
Recall: Plotting $Y = (X - 0.5n)/(0.5\sqrt{n})$

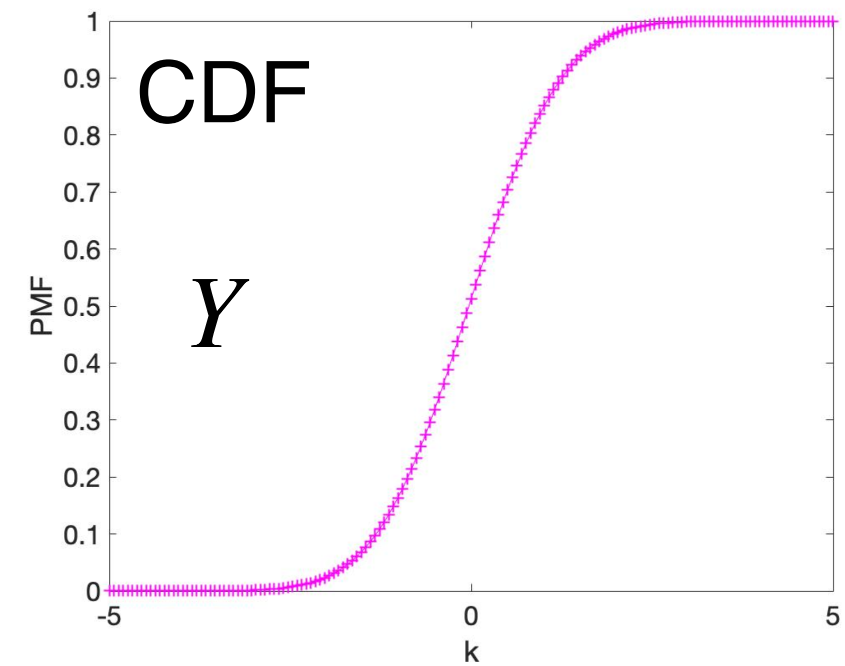
$X \sim \text{Binomial}(n, \frac{1}{2})$ $\quad \frac{1}{2} \cdot n$

$E[X]$ $\quad \sqrt{Var[X]}$

$n = 100$

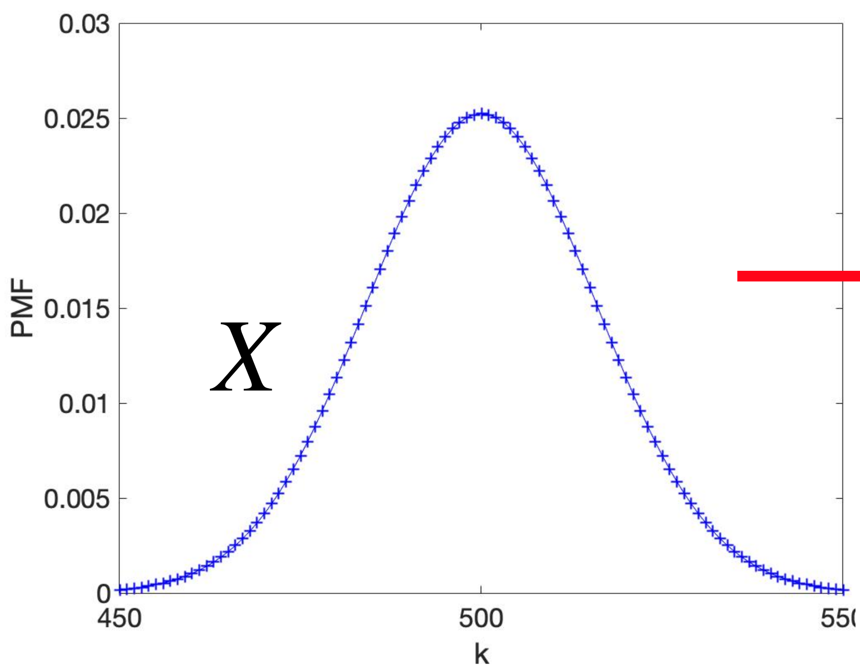


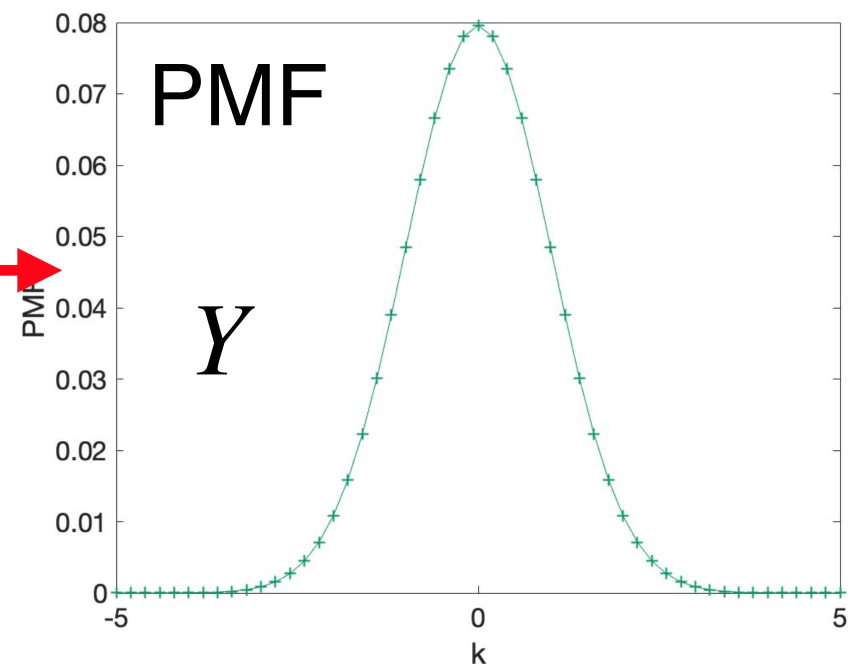


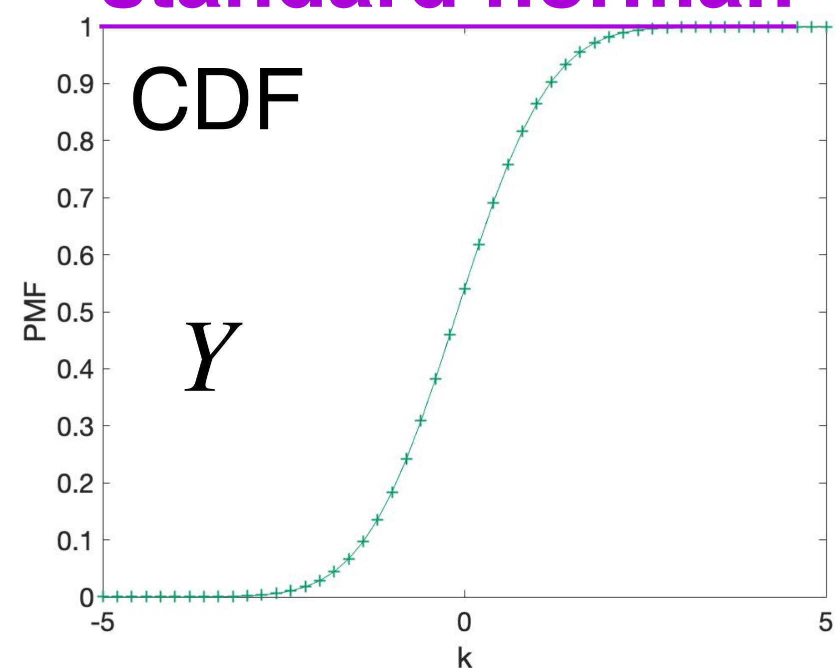


≈ CDF of a standard normal!

$n = 1000$





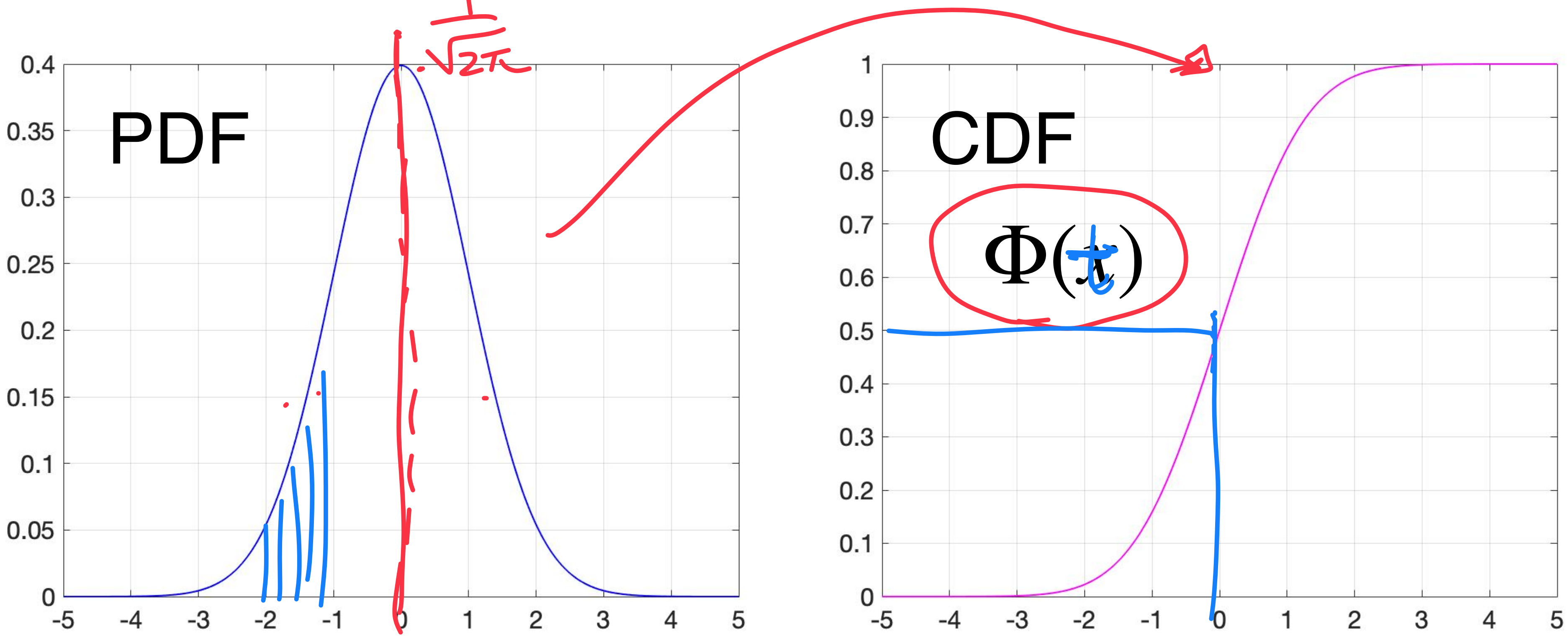


2. Standard Normal Random Variables (Formally)

Standard Normal Random Variables: A random variable X is called standard normal if its PDF is

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

- How to plot the PDF and CDF?



$E[X] = 0$

$\text{Var}[X] = 1$

2. CDF of Standard Normal (Formally)

- As standard normal is widely applicable, we use a special notation $\Phi(\,\cdot\,)$ for its CDF

CDF of Standard Normal: The CDF of a standard normal random variable X is

$$\Phi(t) := P(X \le t) = \int_{-\infty}^{t} \underbrace{\frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right)}_{f_X(x)} dx$$

- **Question**: How to plot $\Phi(t)$?
  - $\underbrace{\Phi(\infty) = ?}_{= P(X \le \infty) = 1}$ $\underbrace{\Phi(0) = ?}_{= P(X \le 0) = 0.5}$

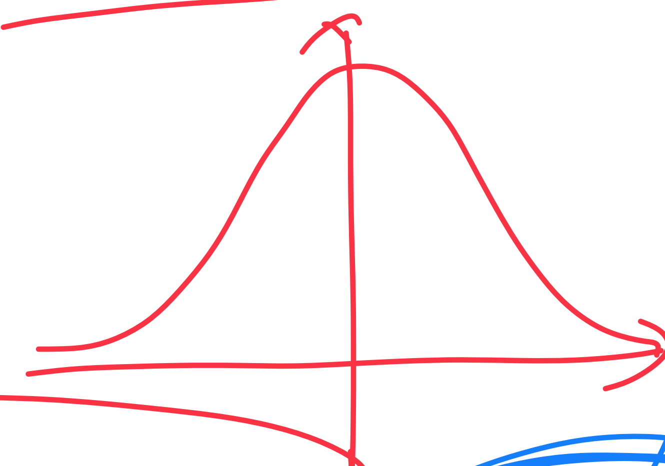
Why is Standard Normal Useful?

1. **Central Limit Theorem**:

$X_1, X_2, X_3, \cdots X_n$

$$Y = \frac{(X_1+X_2+X_3+\cdots X_n) - E[X_1+X_2+\cdots+X_n]}{\sqrt{Var(X_1+X_2+\cdots+X_n)}}$$

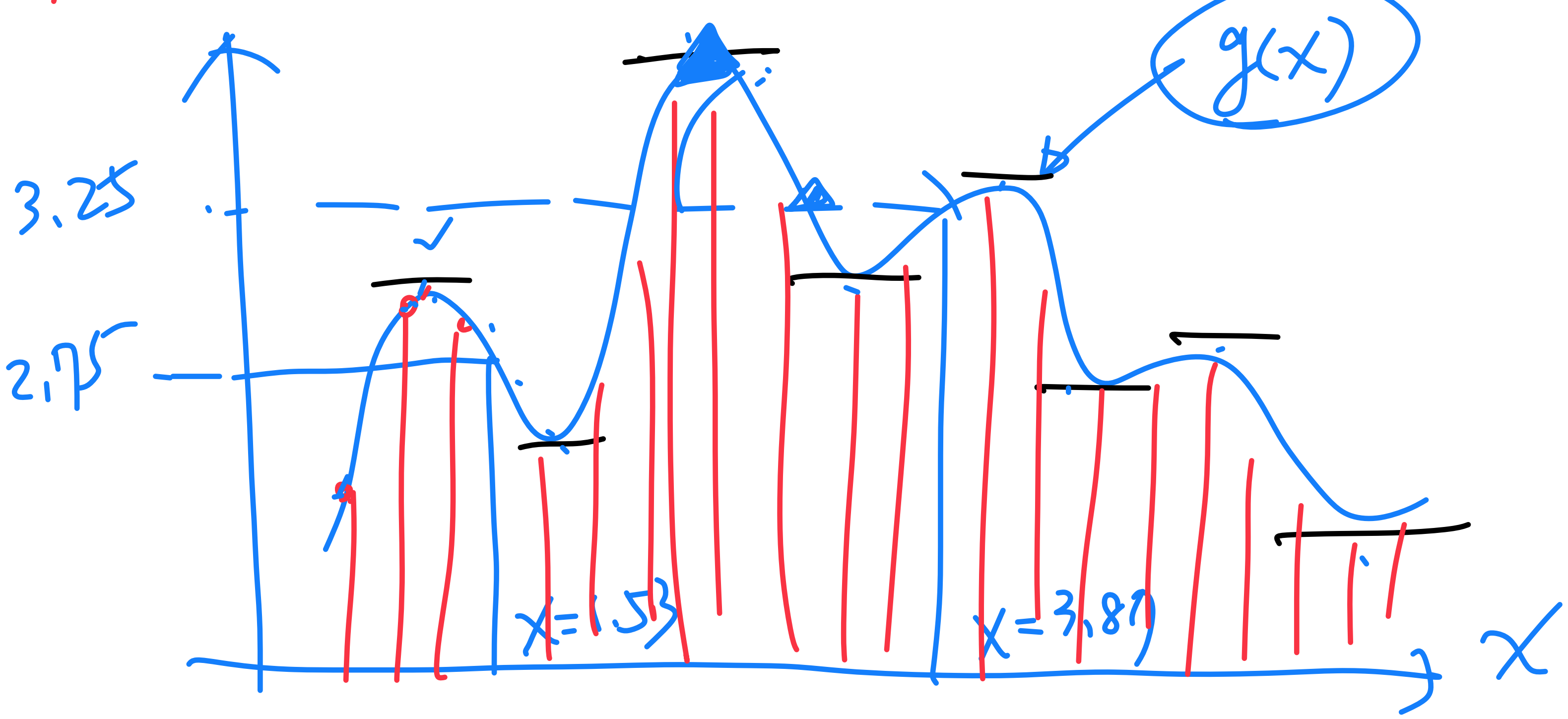
Let $n \to \infty$:

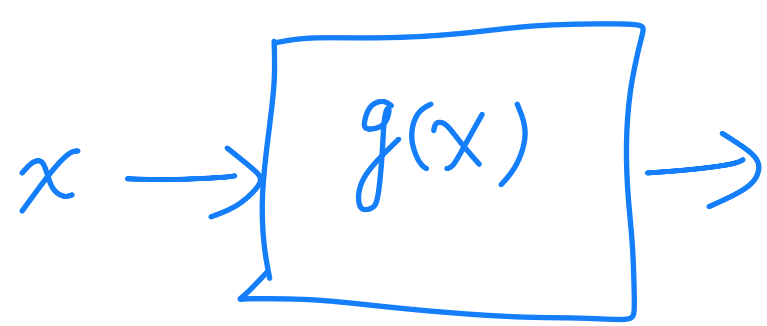


2. **Gaussian Process and Black-Box Optimization**:

multivariate normal

max $g(x)$





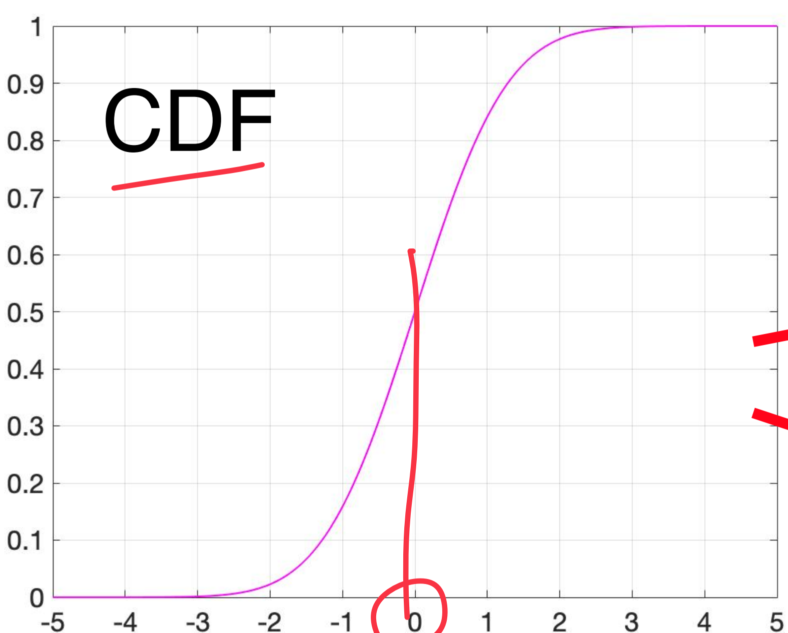
From Standard Normal to Normal: Linear Transformation of Random Variables

X

$Y = 2X + 3$

From Standard Normal to Normal: CDF

- X is standard normal



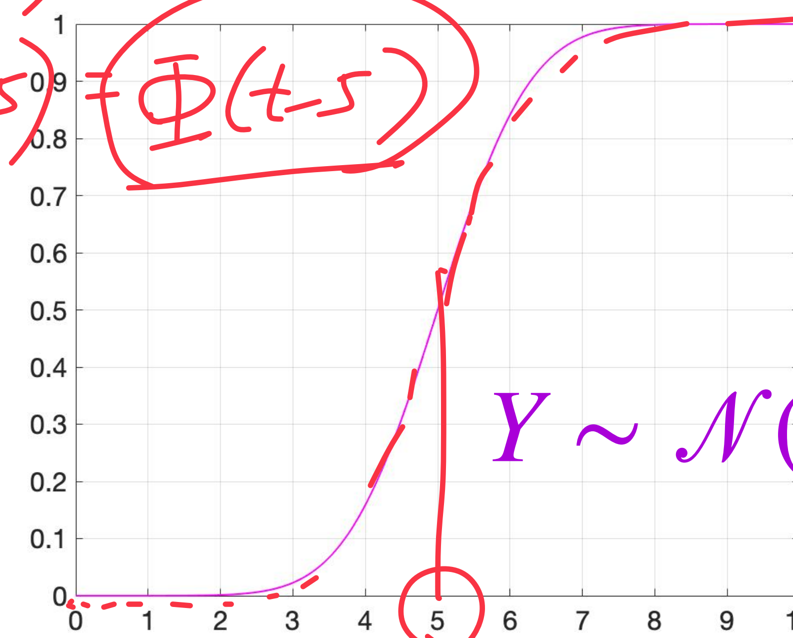
$F_X(t) = \Phi(t)$

$X \sim \mathcal{N}(0,1)$

$Y = X + 5$

$F_Y(t) = P(Y \le t) = P(X+5 \le t) = P(X \le t-5) = \Phi(t-5)$

$F_Y(t) = \Phi(t-5)$

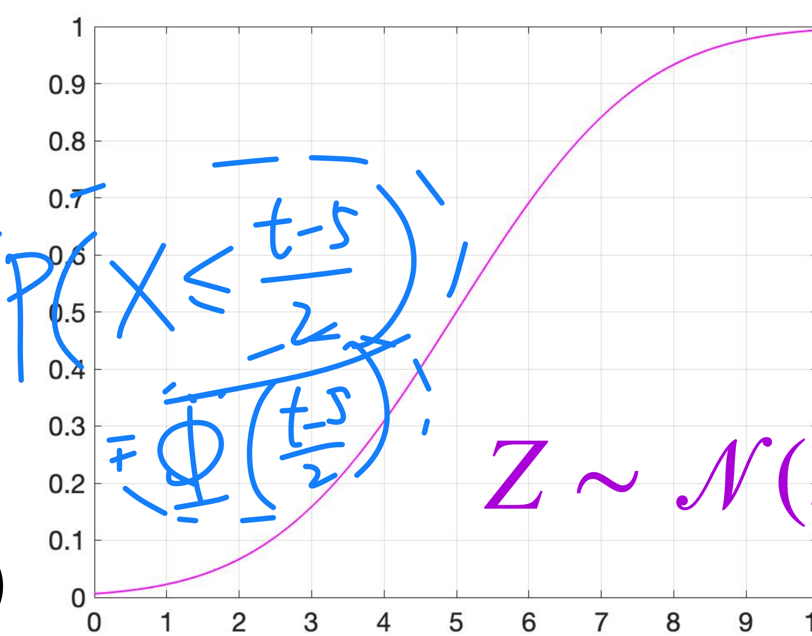


$Y \sim \mathcal{N}(5,1)$

$Z = 2X + 5$

$F_Z(t) = P(Z \le t) = P(2X+5 \le t) = P\left(X \le \frac{t-5}{2}\right) = \Phi\left(\frac{t-5}{2}\right)$

$F_Z(t) = \Phi\left(\frac{t-5}{2}\right)$

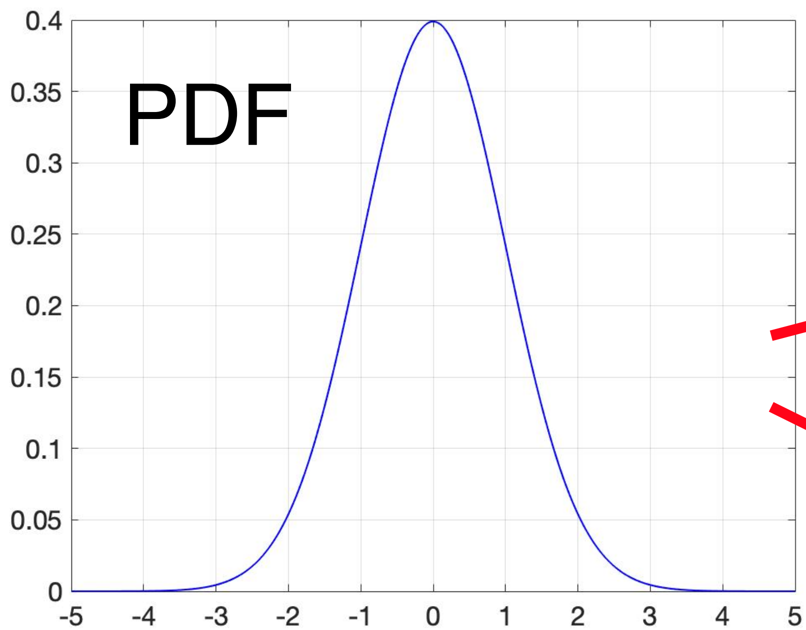


$Z \sim \mathcal{N}(5,2^2)$

From Standard Normal to Normal: PDF

$E[X]$, $Var[X]$

- X is standard normal



$$f_X(t) = \Phi'(t)$$

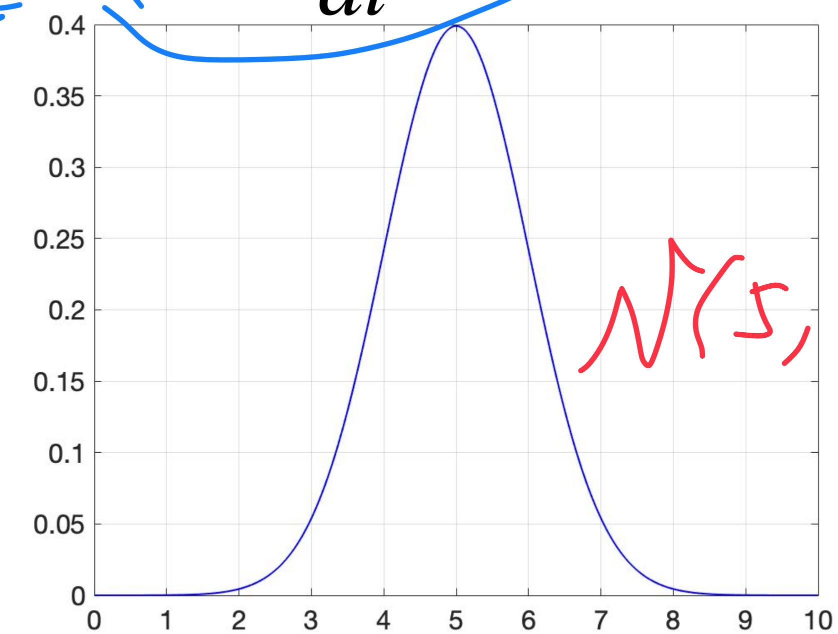
$Y = X + 5$

$E[Y] = 5$

$Var[Y] = 1$

$F_Y(t) = \Phi(t-5)$

$$f_{\cancel{X}\,Y}(t) = \frac{d\Phi(t-5)}{dt} = \Phi'(t-5).$$



$N(5, 1)$

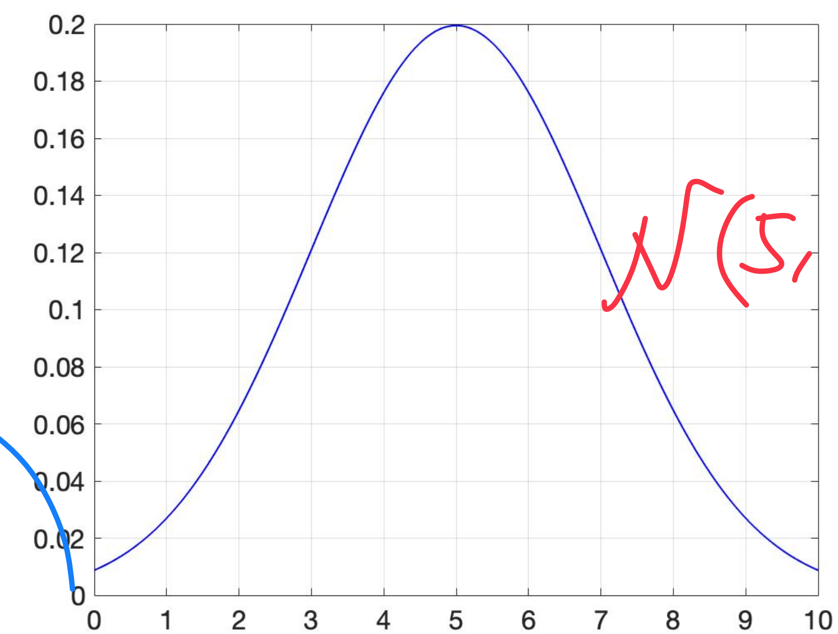
$Z = 2X + 5$

$E[Z] = 5$

$Var[Z] = 4$

$F_Z(t) = \Phi\left(\frac{t-5}{2}\right)$

$$f_{\cancel{X}\,Z}(t) = \frac{d\Phi(\frac{t-5}{2})}{dt} = \frac{1}{2}\Phi'\left(\frac{t-5}{2}\right)$$

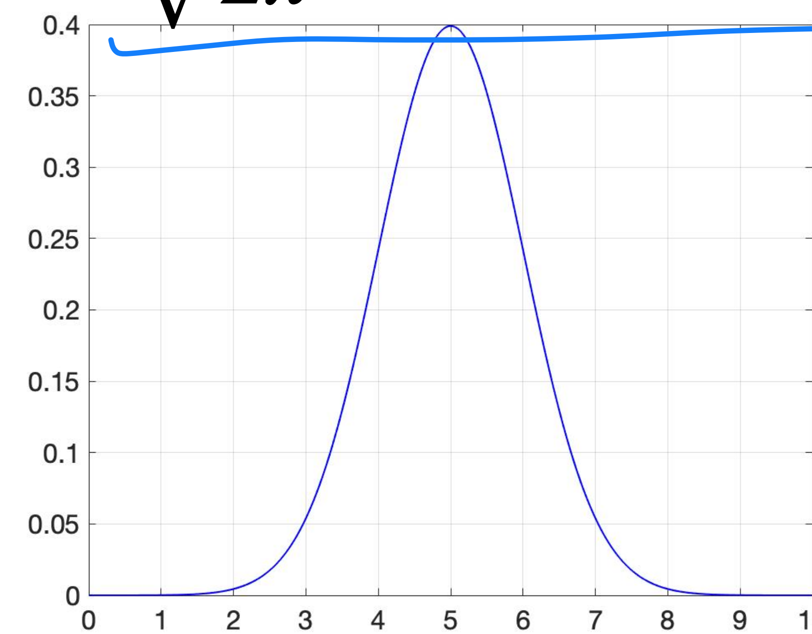


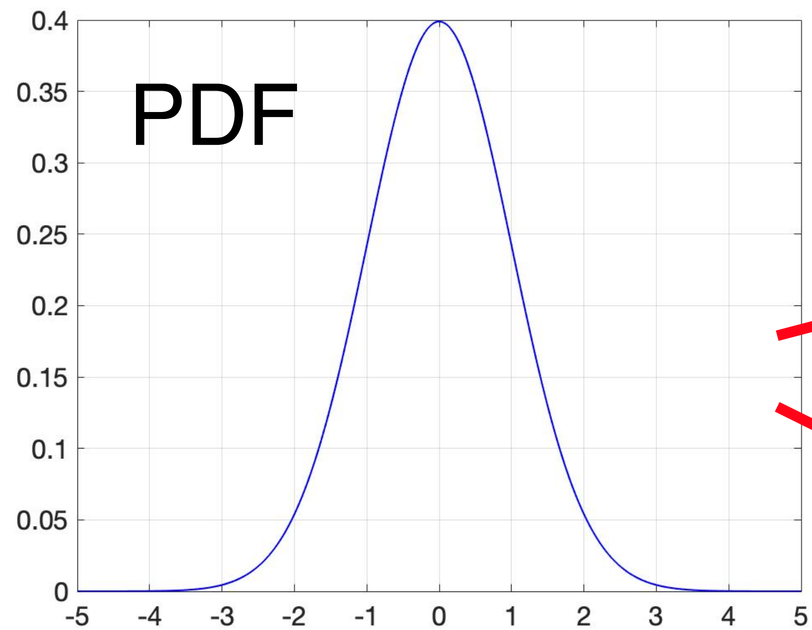
$N(5, 4)$

From Standard Normal to Normal: PDF Φ'

- X is standard normal

$$f_Y(t) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(t-5)^2}{2}\right)$$

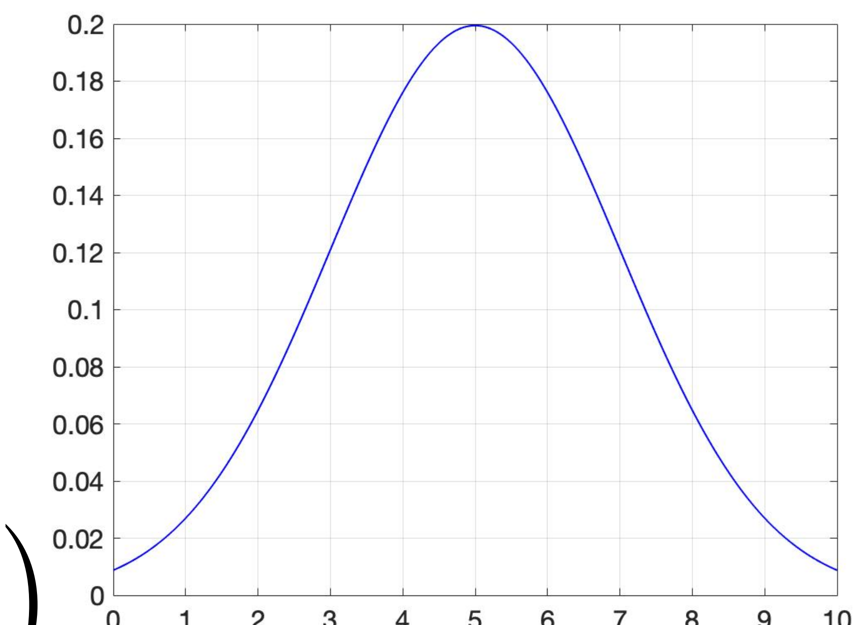




$$Y = X + 5$$

$$Z = 2X + 5$$

$$f_X(t) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-t^2}{2}\right)$$



$$f_Z(t) = \frac{1}{2\sqrt{2\pi}} \exp\left(\frac{-(t-5)^2}{2 \cdot 2^2}\right)$$

A General Recipe for Linear Transformation

- X is a continuous random variable
  - CDF: $F_X(t)$
  - PDF: $f_X(t) = \dfrac{dF_X(t)}{dt}$
- Consider $Y = aX + b$, $a, b \in \mathbb{R}, a \neq 0$
  - CDF $F_Y(t)$?
  - PDF $f_Y(t)$?
  - If $X \sim \mathcal{N}(0,1)$, then $F_Y(t)$ = ?

$F_Y(t) = P(Y \leq t)$

$= P(aX + b \leq t)$

$= P(X \leq \frac{t-b}{a})$

$f_Y(t) = F'_Y(t)$

Normal Random Variables

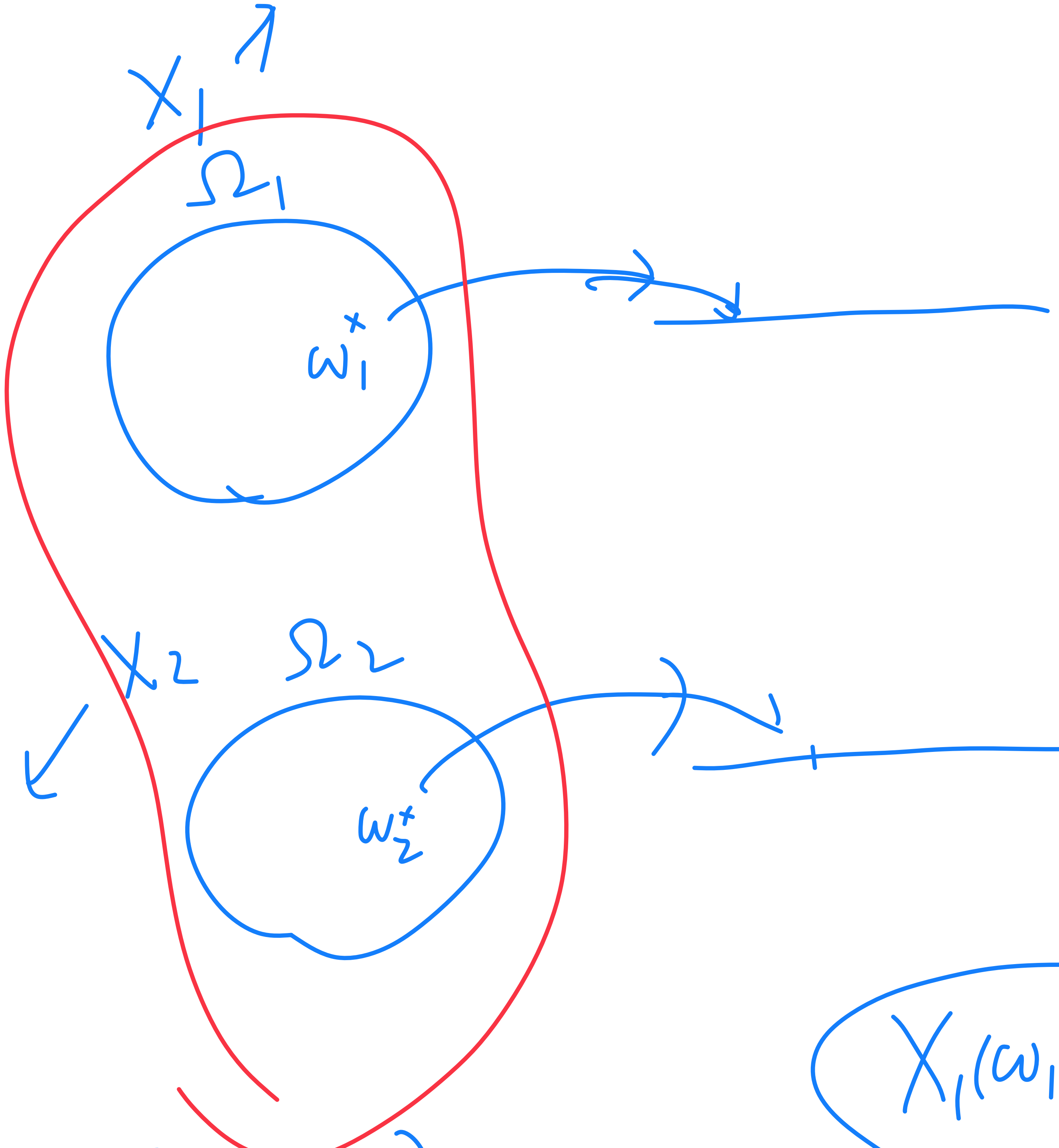
Normal Random Variables: A random variable X is called normal with parameters μ, σ if its PDF is

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right)$$

- Notation: $X \sim \mathcal{N}(\mu, \sigma^2)$

  E[X] Var[X]

- How to plot the PDF?



$$X = \min(X_1, \cdots, X_n)$$

$$\Omega \to \mathbb{R}$$

$$\omega = (\omega_1, \omega_2, \cdots, \omega_n)$$

$$X(\omega) = \min\big(X_1(\omega_1), X_2(\omega_2), \cdots, X_n(\omega_n)\big)$$

$X_1(\omega_1) = 1$ p

$X_2(\omega_2) = 1$ p

$$X(\omega = \{\omega_1, \omega_2\}) = p^2$$

$$X = \min(X_1, X_2)$$

$$P(X = 6) = P(X_1 = 6 \text{ and } X_2 = 6) = P(X_1 = 6) \cdot P(X_2 = 6)$$

Example: Normal Distribution

- Example: Let $X \sim \mathcal{N}(-2,5)$
  - What is $P(|X| < 4)$?

Exponential Random Variables

Recall: Geometric Random Variables

- Suppose $X \sim$ Geometric(p)
  - What is the PMF of X?
  - Memoryless property?

- Question: Is there a continuous counterpart of a geometric random variable?

3. Exponential Random Variables

Exponential Random Variables: A random variable X is exponential with parameters $\lambda > 0$ if its PDF is

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

- How to plot the PDF of Exp($\lambda = 1$)?

3. Exponential Random Variables

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

- What is the CDF of X?

Memoryless Property

- Suppose $X \sim \text{Exp}(\lambda)$
  - What is $P(X > s + t \,|\, X > t)$?

Example: Nokia 3310

- Example: Suppose the lifetime of a Nokia 3310 is an <u>exponential</u> random variable with <u>mean = 10 years</u>.
  - Suppose a Nokia 3310 was bought 15 years ago.
  - $P($it will last another 5-10 years$)$?

Exponential Distribution: A Good Model for Occurrence of Events

- Communication networks: Inter-arrival time between two data packets
- Survival analysis: User’s lifetime (App, social network…)
- Reliability modeling: Amount of time until the hardware on AWS EC2 fails

1-Minute Summary

1. Special Continuous Random Variables

- Standard Normal and Normal
- Exponential and Memoryless Property