LECTURE #3
PROBABILITY REVIEW (II)

204528
Queueing Theory and Applications in Networks

Assoc. Prof. Anan Phonphoem, Ph.D. (รศ.ดร. อนันต์ ผลเพิ่ม)
Computer Engineering Department, Kasetsart University

Outline

- Probability meaning
- Random Variable
- Discrete RV
- Continuous RV
- Multiple RVs

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Cumulative Distribution Function (CDF)

- Definition:

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

- Contain complete information about the probability model of the random variable
- PMF ⟺ CDF

CDF Theorem

Theorem: For a discrete random variable X

with $S_X=\{x_1,x_2,\ldots\}$ & $\mathbf{x_1 \leq x_2 \leq \ldots}$

1) $F_x(-\infty) = 0$ and $F_x(\infty) = 1$ → **From 0 to 1**

2) $\forall x' \geq x,\ F_x(x') \geq F_x(x)$ → **Monotonic Increasing**

3) For $x_i \in S_x$ and ε = +small number

$F_x(x_i) - F_x(x_i - \varepsilon) = P_x(x_i)$ → **Discontinuity = $P_X(x)$**

4) $F_x(x) = F_x(x_i) \quad \forall x,\ x_i \leq x < x_{i+1}$ → **Horizon line**

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

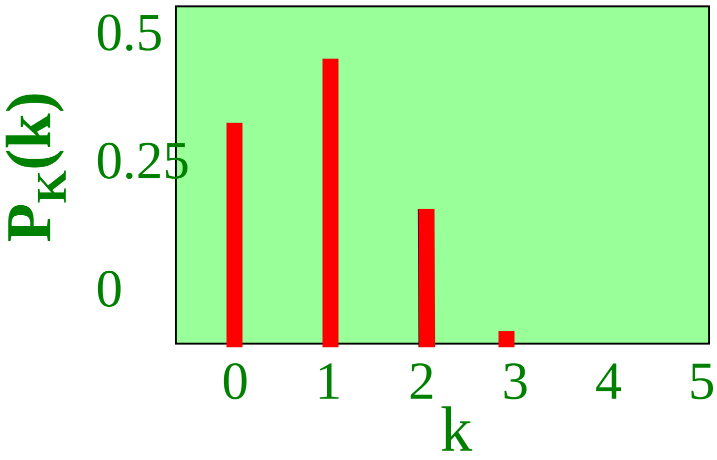
CDF Example

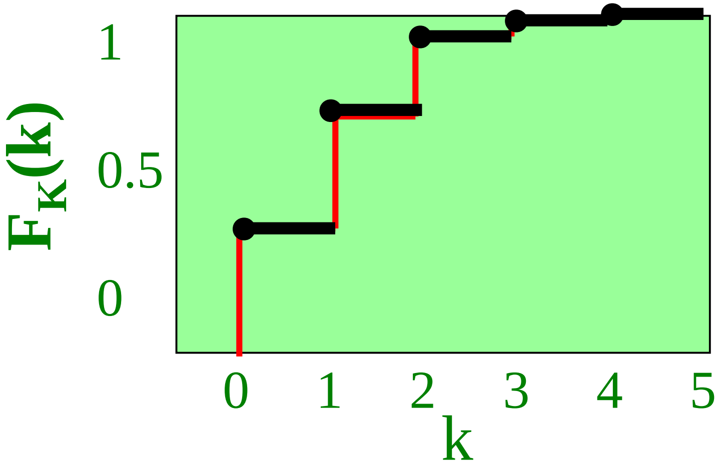
- For a binomial RV, # of fail CDs in 5 tests with p = 0.2

$$P_K(k) = \begin{cases} \binom{5}{k} (0.2)^k (0.8)^{5-k} & k = 0,1,2,\ldots,5 \\ 0 & \text{Otherwise} \end{cases}$$

CDF Example

k	$P_K(k)$	k	$P_K(k)$
0	0.33	3	0.05
1	0.41	4	0.01
2	0.20	5	0





Average

- Study RV → average
- What is the average of an RV?
  - A single number that describes the RV
  - An example of statistic
- What is Statistic?
  - Numbers that collect all information of things under our interesting
  - Averages: mean, mode, and median

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Average

- Mean:
  - Sum / #terms
- Mode:
  - Most common value
  - $P_X(x_{mod}) \geq P_X(x) \quad \forall x$
- Median:
  - The middle of the data set
  - $P[X < x_{med}] = P[X > x_{med}]$

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Mean → Expected Value

- Adding all measurements / #terms

$$E[X] = \mu_X = \sum_{x \in S_X} x P_X(x)$$

- Example:

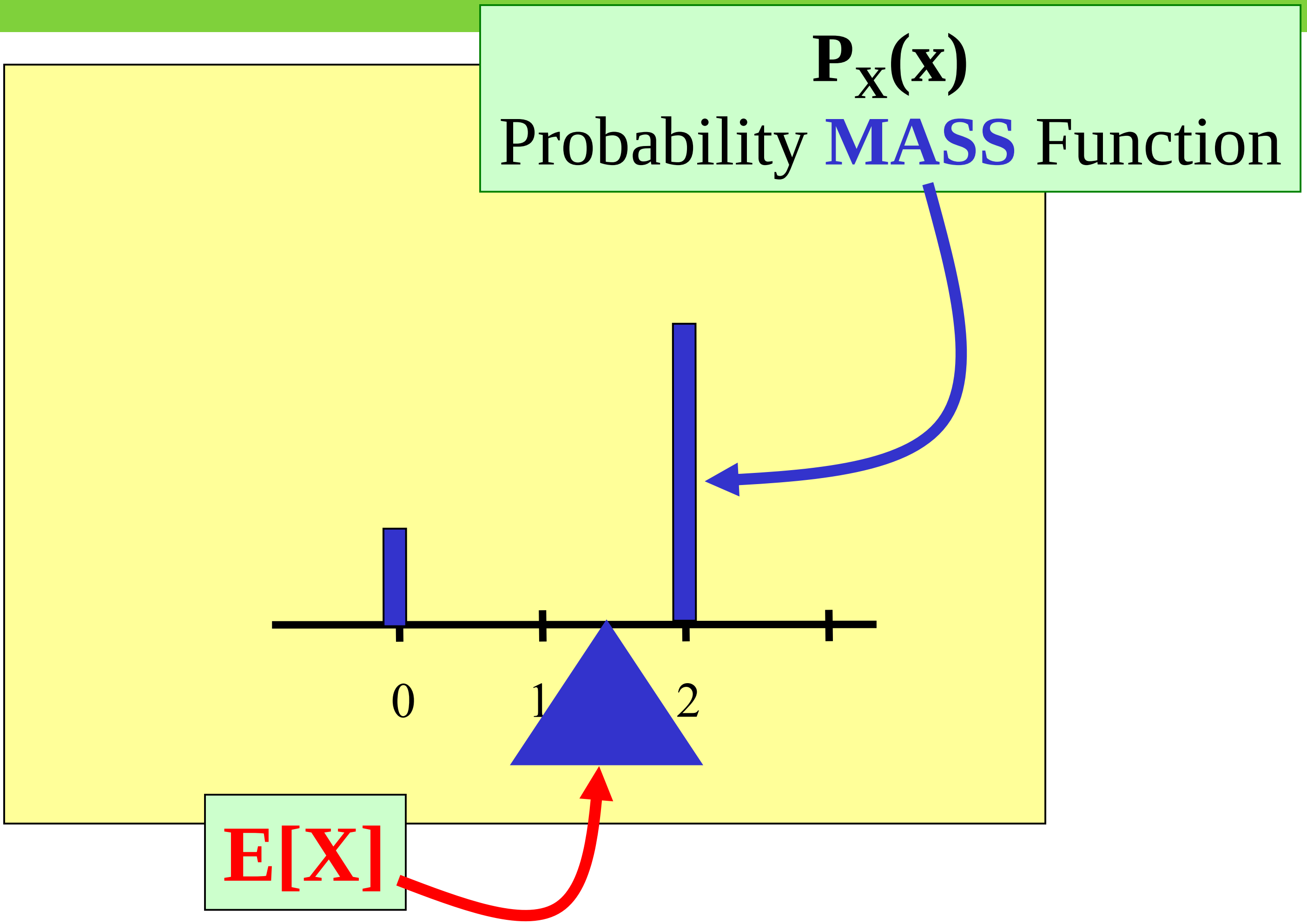
$$P_T(t) = \begin{cases} 1/4 & t = 0 \\ 3/4 & t = 2 \\ 0 & \text{Otherwise} \end{cases}$$

- $E[T] = ?$

$$= 0(1/4) + 2(3/4) = 3/2$$

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Expected Value



Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Useful Discrete RV

Uniform Equiprobable outcomes	$\begin{cases} 1/(j-k+1) & x = k,k+1,k+2,..,j \\ 0 & \text{Otherwise} \end{cases}$	$E[X] = \frac{(j+k)}{2}$
Bernoulli Pass/Fail	$\begin{cases} 1-p & x = 0 \\ p & x = 1 \\ 0 & \text{Otherwise} \end{cases}$	$E[X] = p$
Geometric # tests until fail	$\begin{cases} p(1-p)^{x-1} & x = 1,2,3,\ldots \\ 0 & \text{Otherwise} \end{cases}$	$E[X] = 1/p$

Useful Discrete RV

Binomial # fails in n tests	$\begin{cases} \binom{n}{x} p^x(1-p)^{n-x} & x=1,2,\ldots,n \\ 0 & \text{Otherwise} \end{cases}$	$E[X] = np$
Pascal # tests until k fails	$\begin{cases} \binom{x-1}{k-1} p^k(1-p)^{x-k} & x = k,k+1,\ldots \\ 0 & \text{Otherwise} \end{cases}$	$E[X] = k/p$
Poisson occurrence in a period	$\begin{cases} \dfrac{(\lambda T)^x e^{-(\lambda T)}}{x!} & x = 0,1,2\ldots \\ 0 & \text{Otherwise} \end{cases}$	$E[X] = \alpha$ $\alpha = \lambda T$

Variance & Standard Deviation

- We knew average, E[X],
  why do we need these Variance & Standard Deviation?
- How far from the average?
- $T = X - \mu_X$

  $E[T] = E[X - \mu_X]$

  $= 0$

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Variance

- The useful measurement is **E[|T|]**
- $E[T^2] = E[(X - \mu_x)^2]$ → **Variance**

Definition:

$$Var[X] = E[(X - \mu_x)^2]$$

Standard Deviation

Definition:

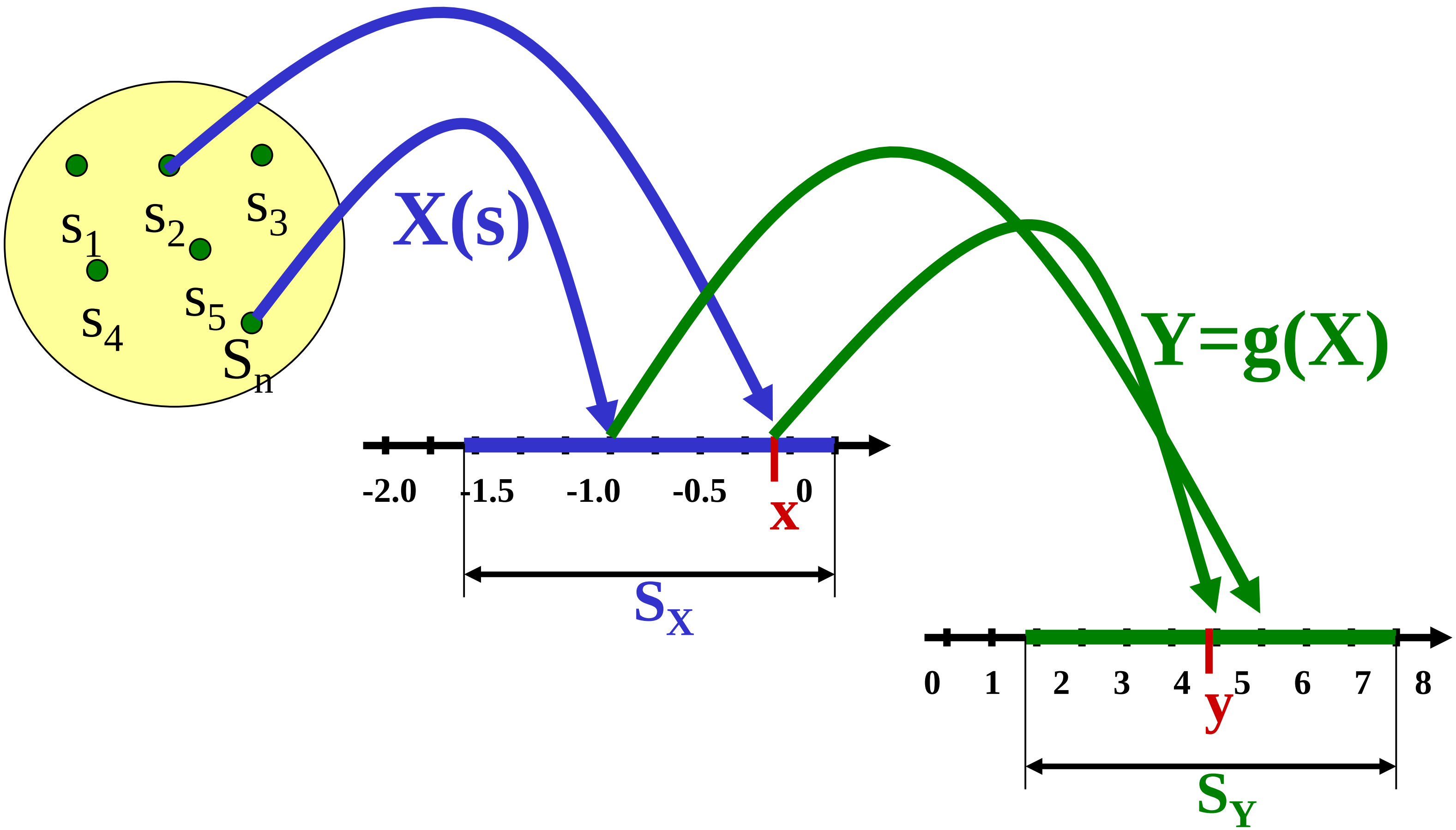
Sigma X

$$\sigma_X = \sqrt{Var[X]}$$

- σ_X can compare to μ_x
- Ex. σ_X = 15, Score +6 from mean
  → OK. Middle of class
- Ex. σ_X = 3,Score +6 from mean
  → V.Good In Top class group

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Derived Random Variable



Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Why do we need a Derived Random Variable?

- From sample values of the random variable, use these values to compute other quantities.
- Example:
  - Find a decibel value form signal-to-noise ratio
- $Y = g(X)$

Example-1

- Random Variable X = # pages in one fax
- $P_X(x)$ = number of pages in each fax
- Charging plan
  - 1^{st} page = 10 Baht
  - 2^{nd} page = 9 Baht
  - …
  - 5^{th} page = 6 Baht
  - 6 – 10 pages = 50 Baht
- Find the charge in Baht for sending one fax

Example-1

- Random Variable Y = the charge in Baht for sending one fax

$$Y = g(X) = \begin{cases} 10.5X - 0.5X^2 & 1 \le X \le 5 \\ 50 & 6 \le X \le 10 \end{cases}$$

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

PMF of Y

Theorem:

$$P_Y(y) = \sum_{x:g(x)=y} P_X(x)$$

P[Y=y] = Σ of all outcomes X = x for which Y = y

Conditional PMF

$$P[A|B] = P[X = x|B]$$

Definition: Given event B, P[B] > 0

$$\mathbf{P_{X|B}(x) = P[X=x|B]}$$

Theorem: $P[A] = \sum_{i=1}^{n} P[A|B_i]P[B_i]$

Theorem: $P_X(x) = \sum_{i=1}^{n} P_{X|B_i}(x)P[B_i]$

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Outline

- Probability meaning
- Random Variable
- Discrete RV
- Continuous RV
- Multiple RVs

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

MULTIPLE DISCRETE RVS

What is Multiple Discrete RV?

- Each observation → Random Variable
- 2 observations → 2 Random Variables
- ≥ 2 observations → Multiple RVs

Joint PMF

- For an experiment, Observe one thing
  - Model with one Random Variable
  - Describe the prob. model by using PMF
- For the same experiment, Observe 2 things
  - 2 Random Variables → X and Y
  - Joint PMF
- $P_{X,Y}(x,y)$

Joint PMF

Definition:

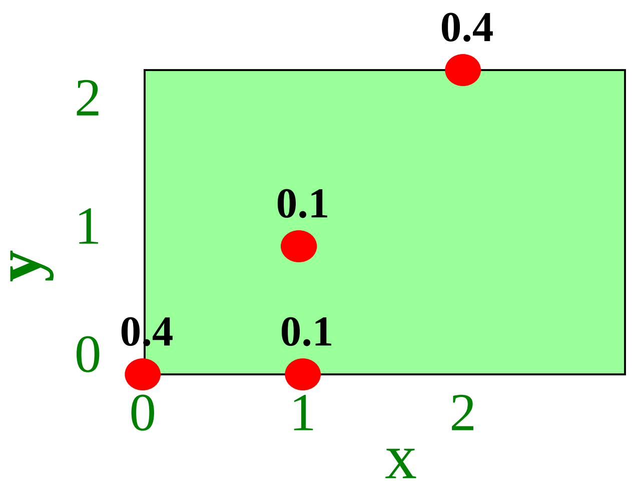
$$P_{\mathbf{X,Y}}(x,y) = P[X=x,\ Y=y]$$

$$S_{X,Y} = \{(x,y) \mid P_{X,Y}(x,y) > 0\ \}$$

3 forms of Joint PMF

$$P_{X,Y}(x,y) = \begin{cases} 0.4 & x=2, y=2 \\ 0.1 & x=1, y=1 \\ 0.1 & x=1, y=0 \\ 0.4 & x=0, y=0 \\ 0 & \text{Otherwise} \end{cases}$$

$P_{X,Y}(x,y)$	y=0	y=1	y=2
x=0	0.4	0	0
x=1	0.1	0.1	0
x=2	0	0	0.4



Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Marginal PMF

- In an experiment with 2 RVs, X and Y
  - Possible to consider only one (X) and ignore Y
  - $P_X(x)$

Theorem: For random variables X and Y with joint PMF $P_{X,Y}(x,y)$:

$$\mathbf{P_X(x) = \sum_{y} P_{X,Y}(x,y)}$$

$$\mathbf{P_Y(y) = \sum_{x} P_{X,Y}(x,y)}$$

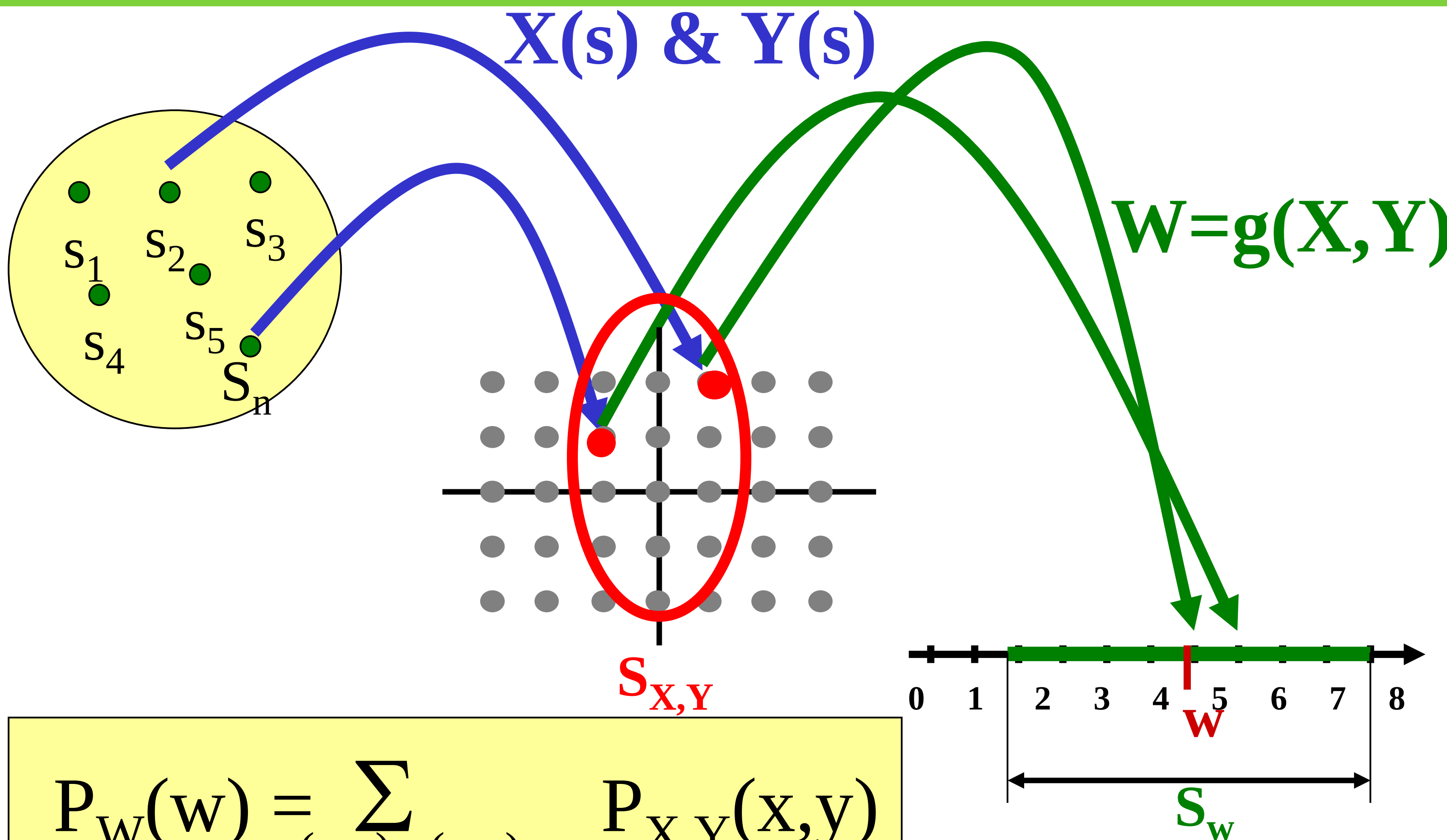
Marginal PMF

Margin

$P_{X,Y}(x,y)$	y=0	y=1	y=2	$P_X(x)$
x=0	0.4	0	0	0.4
x=1	0.1	0.1	0	0.2
x=2	0	0	0.4	0.4
$P_Y(y)$	0.5	0.1	0.4	1

Margin

Derived Random Variable
Functions of 2 RVs



$$P_W(w) = \sum_{(x,y):g(x,y)=w} P_{X,Y}(x,y)$$

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Expected Value of g(X,Y)

Theorem: for $W = g(X,Y)$

$$E[W] = \sum_{x \in S_X} \sum_{y \in S_Y} g(X,Y)\ P_{X,Y}(x,y)$$

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

For any 2 RVs

Theorem:

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

- Find E[X] and E[Y]
  → Marginal PMF

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Var[X+Y]

Definition: $Var[X] = E[(X - \mu_X)^2]$

$$Var[X+Y] = E[((X+Y) - \mu_{X+Y})^2]$$
$$= E[((X+Y) - (\mu_X + \mu_Y))^2]$$
$$= E[((X-\mu_X) + (Y-\mu_Y))^2]$$
$$= E[(X-\mu_X)^2 + 2(X-\mu_X)(Y-\mu_Y) + (Y-\mu_Y)^2]$$
$$= E[(X-\mu_X)^2] + 2E[(X-\mu_X)(Y-\mu_Y)] + E(Y-\mu_Y)^2]$$

Theorem:

$$Var[X+Y] = Var[X] + Var[Y] + 2\,E[(X-\mu_X)(Y-\mu_Y)]$$

Covariance

Covariance of X and Y

Definition: $Cov[X,Y] = E[(X-\mu_x)(Y-\mu_Y)]$

$$
\begin{aligned}
Cov[X,Y] &= E[XY - \mu_x Y - \mu_Y X + \mu_x \mu_y] \\
&= E[XY] - \mu_x E[Y] - \mu_Y E[X] + \mu_x \mu_y \\
&= E[XY] - \mu_x \mu_y - \mu_x \mu_y + \mu_x \mu_y
\end{aligned}
$$

Theorem: $Cov[X,Y] = E[XY] - \mu_x \mu_y$

Covariance of X and Y

Theorem: $Cov[X,Y] = E[XY] - \mu_x \mu_y$

Correlation

If $X = Y$ → $Cov[X,X] = E[XX] - \mu_x \mu_x$

$= E[X^2] - \mu_x^{\,2}$

$= E[X^2 - 2\mu_x^2 + \mu_x^2]$

$= E[X^2 - 2\mu_x X + \mu_x^2]$

$= E[(X - \mu_x)^2]$

$= Var[X]$

If μ_x or $\mu_Y = 0$ → $Cov[X,Y] = E[XY]$

Correlation

Definition: The correlation of X and Y is $r_{X,Y}$

$$r_{X,Y} = E[XY]$$

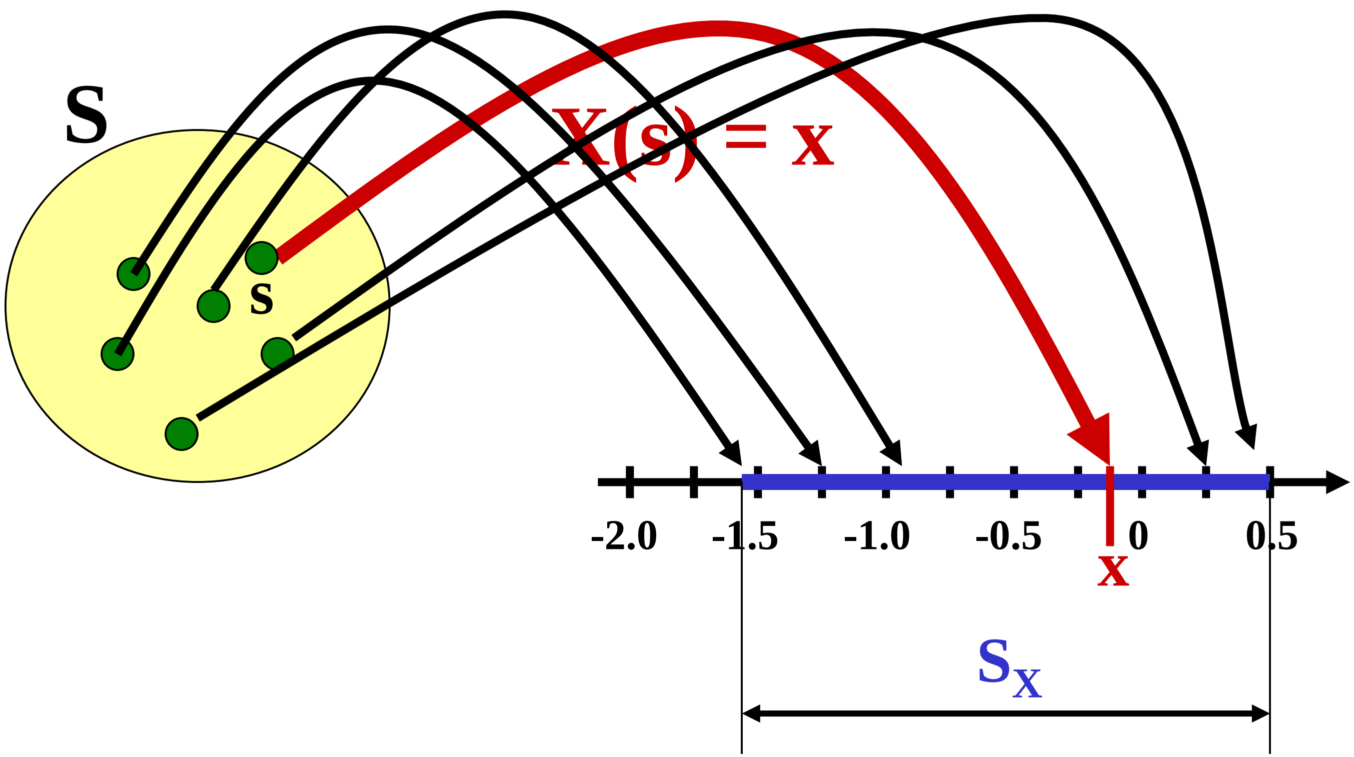
Theorem: $\mathrm{Cov}[X,Y] = r_{X,Y} - \mu_x \mu_y$

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

CONTINUOUS RANDOM VARIABLE

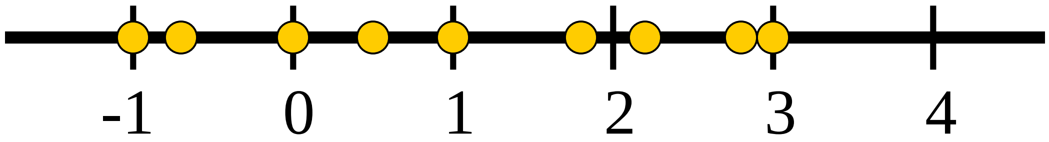
Random Variable

X is a function that maps each outcome, s, in S to a real number X(s), x



Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Continuous Sample Space



In **Discrete**: countable set of numbers

$S_X = \{-1,0,1,3,4\}$

$S_Y = \{-1,-0.9,0,0.5,1,1.8,2.25,2.9,3\}$

In **Continuous**: uncountable set of numbers

S_X = **Interval** between 2 limits

$S_X = (x_1,x_2) = (-1,3)$

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Probability of a continuous RV

- Measuring T, the download time

$$S_T = \{t \mid 0 < t < 12\}$$

- Guess the download time is (0, 10] minutes
- Guess the download time is [5, 8] minutes
- Guess the download time is [5, 5.5] minutes

Chance that our guess is correct is decreasing

- Guess the download time is exactly 5.25 min.

Probability of each individual outcome is zero.
The interesting probability is an **interval.**

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

CDF

- In discrete:
  - Probability Mass Function (PMF), $P_X(x)$
- In continuous:
  - Impossible to define PMF
  - Cumulative Distribution Function (CDF)

Definition: $F_X(x) = P[X \leq x]$

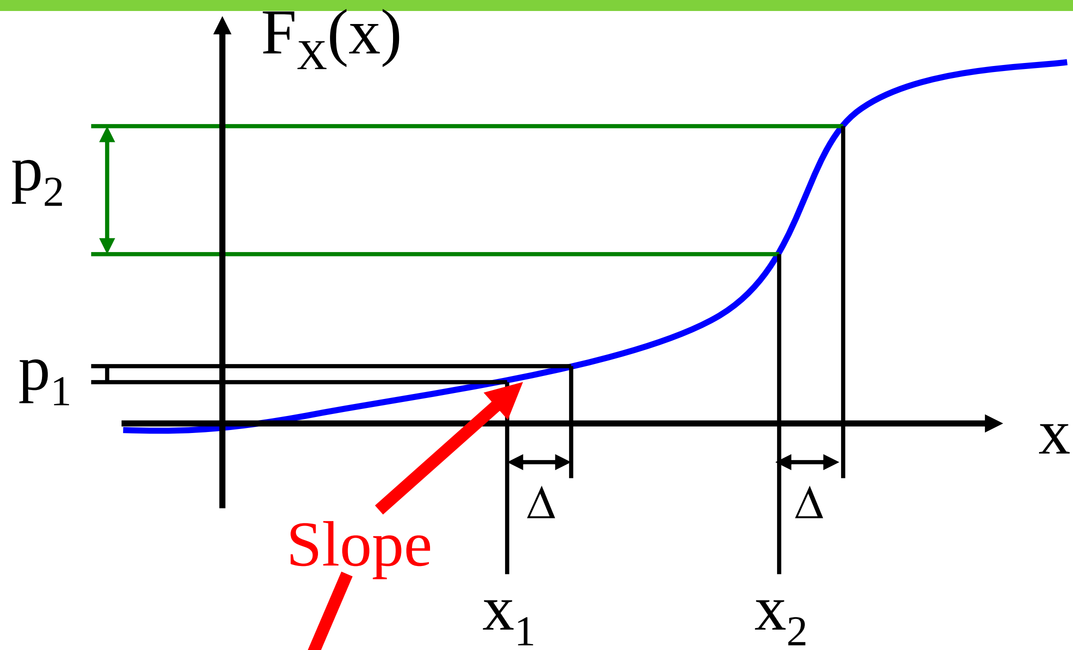
- PMF ⟺ CDF

CDF Theorem

Theorem:

- $F_X(-\infty) = 0$
- $F_X(\infty) = 1$
- $P[x_1 < X \leq x_2] = F_X(x_2) - F_X(x_1)$

Probability Density Function



$$p_1 = P[x_1 < X \leq x_1 + \Delta]$$
$$= F_X(x_1 + \Delta) - F_X(x_1)$$
$$= \frac{F_X(x_1 + \Delta) - F_X(x_1)}{\Delta}\,\Delta$$

$$p_2 = P[x_2 < X \leq x_2 + \Delta]$$
$$= F_X(x_2 + \Delta) - F_X(x_2)$$

For $\Delta \rightarrow 0$,
Slope $\rightarrow dF_X(x)/dx$ at x_1

Probability Density Function

- The slope of CDF in a region near x
  → Probability of random variable X near x
  → The prob. in a small region(Δ) = slope * Δ
- Slope of CDF → PDF

Definition:

Probability Density Function (PDF) is

$$f_X(x) = \frac{dF_X(x)}{dx}$$

PDF Theorem

Theorem:

- $f_X(x) \geq 0 \quad$ for all x
- $F_X(x) = \int_{-\infty}^{x} f_X(u)\, du$
- $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$

Expected Values

For Discrete Random Variable:

$$E[X] = \sum_{x \in S_X} x \, P_X(x)$$

For Continuous Random Variable:

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

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Expected Value & Variance

- Find E[X]

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

- Find E[X²]

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

- Find Var[X]

$$Var[X] = \int_{-\infty}^{\infty} (x - \mu_X)^2\, f_X(x)\, dx$$

Some Useful Continuous RVs

- Uniform
- Exponential
- Gaussian

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Uniform Continuous RV

Definition:

$$f_X(x) = \begin{cases} 1/(b-a) & a \le x < b \\ 0 & \text{Otherwise} \end{cases}$$

where $b > a$

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Uniform Continuous RV

Theorem:

- $$F_X(x) = \begin{cases} 0 & x \le a \\ (x - a)/(b - a) & a < x \le b \\ 1 & x > b \end{cases}$$
- $E[X] = (b + a)/2$
- $Var[X] = (b - a)^2/12$

Exponential Continuous RV

Definition:

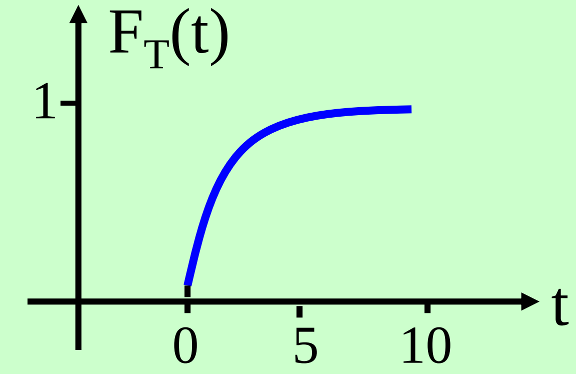
$$f_X(x) = \begin{cases} a\,e^{-ax} & x \geq 0 \\ 0 & \text{Otherwise} \end{cases}$$

where $a > 0$

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Exponential Example

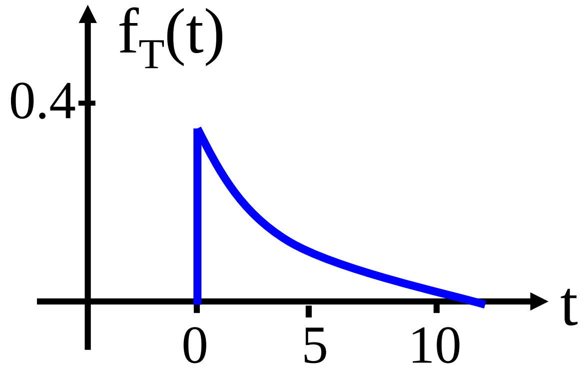
$$F_T(t) = \begin{cases} 1 - e^{-t/3} & t \geq 0 \\ 0 & \text{Otherwise} \end{cases}$$



Find PDF

$$f_T(t) = \frac{dF_T(t)}{dt}$$

$$= \begin{cases} (1/3)\, e^{-t/3} & t \geq 0 \\ 0 & \text{Otherwise} \end{cases}$$



Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Exponential Continuous RV

Theorem:

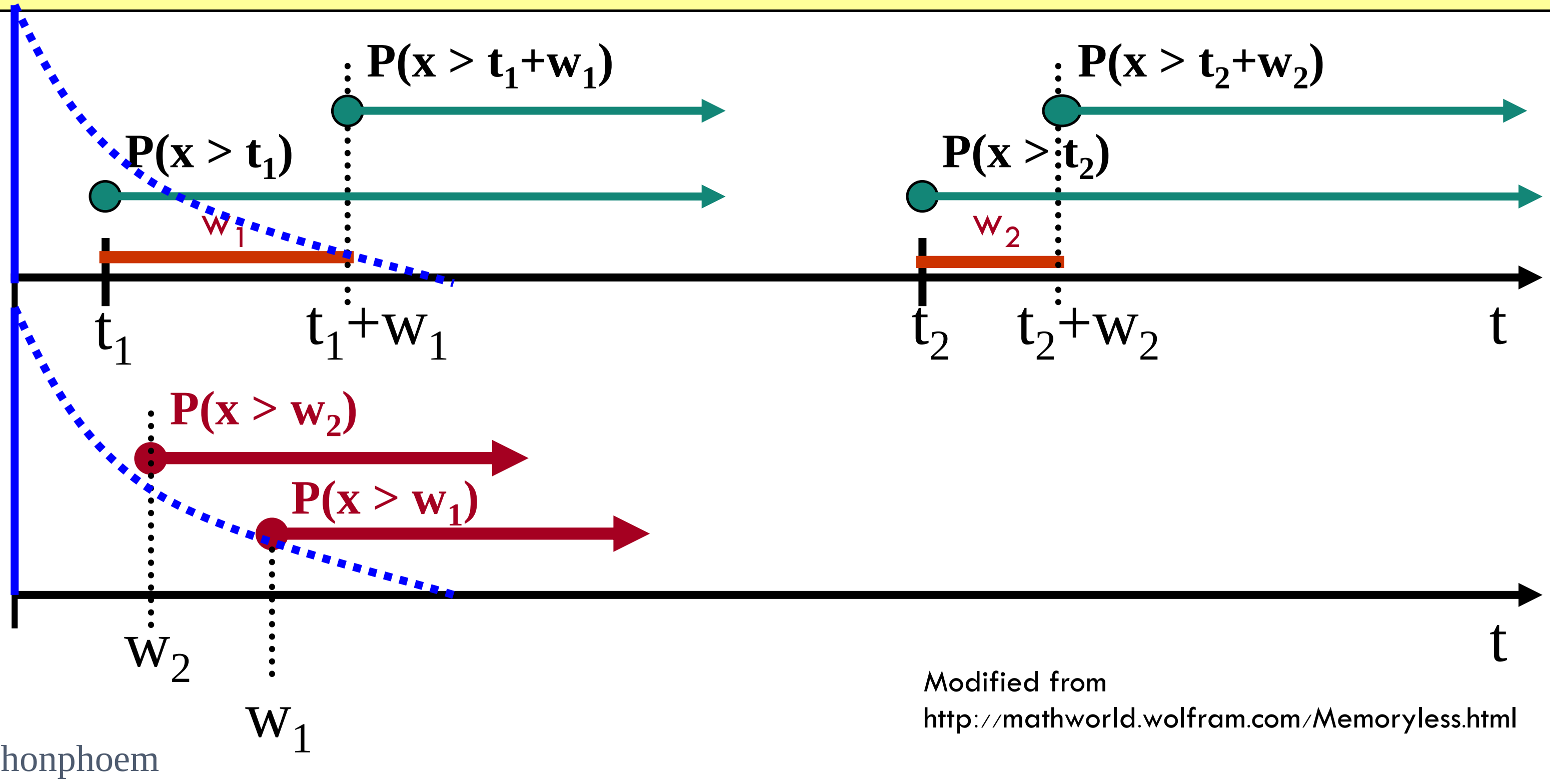
- $$F_X(x) = \begin{cases} 1 - e^{-ax} & x \geq 0 \\ 0 & \text{Otherwise} \end{cases}$$
- $E[X] = 1/a$
- $Var[X] = 1/a^2$

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Memoryless Property

A variable x is memoryless with respect to t if,

$$P(x > t+w \mid x > t) = P(x > w) \quad \forall w \text{ with } t \neq 0$$



Modified from
http://mathworld.wolfram.com/Memoryless.html

Memoryless Property

$$P(x > t+w \mid x > t) = P(x > w)$$

$$\frac{P(x > t+w,\ x > t)}{P(x > t)} = P(x > w)$$

$$P(x > t+w,\ x > t) = P(x > w)\ P(x > t)$$

$$P(x > t+w) = P(x > w)\ P(x > t)$$

Memoryless Property
(Exponential Distribution)

- For Exponential Distribution

$$P(x > t) = e^{-\lambda t}$$

$$P(x > t+w) = e^{-\lambda(t+w)}$$

- Therefore,

$$P(x > t+w) = P(x > w)\,P(x > t)$$

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

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

Exponential distribution is the only **Memoryless** random distribution

If t and w are integers, then the **Geometric distribution** is **Memoryless**

Gaussian Random Variables

Definition:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \, e^{\frac{-(x - \mu)^2}{2\sigma^2}}$$

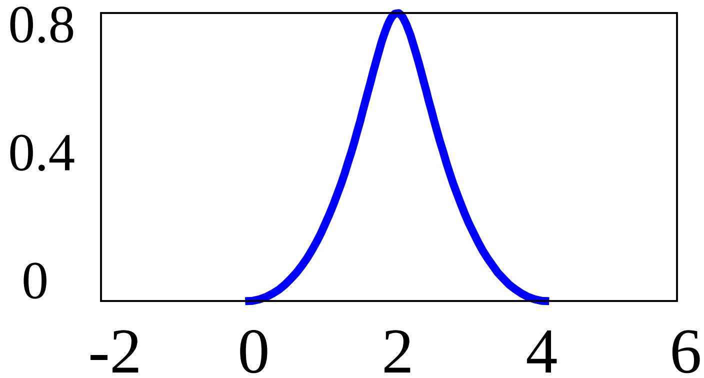
where $\mu \in$ Real, and $\sigma > 0$

- Gaussian RV → **Normal** RV

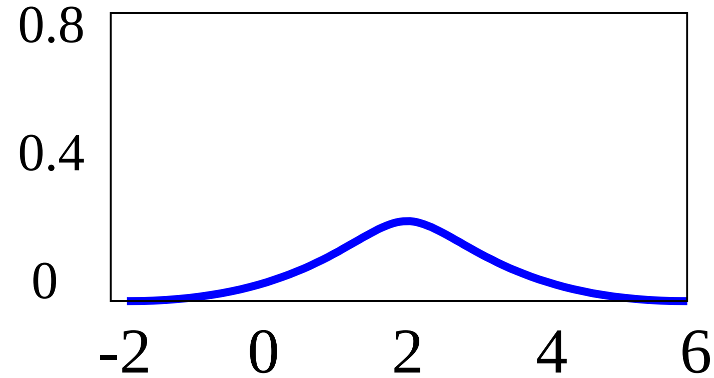
$N[\mu, \sigma^2]$

N → Normal, μ → Mean, σ^2 → Variance

Gaussian Random Variables



$\mu = 2, \ \sigma = 1/2$



$\mu = 2, \ \sigma = 2$

$f_X(x)$ → Bell Shape with 2 parameters: μ and σ

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

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

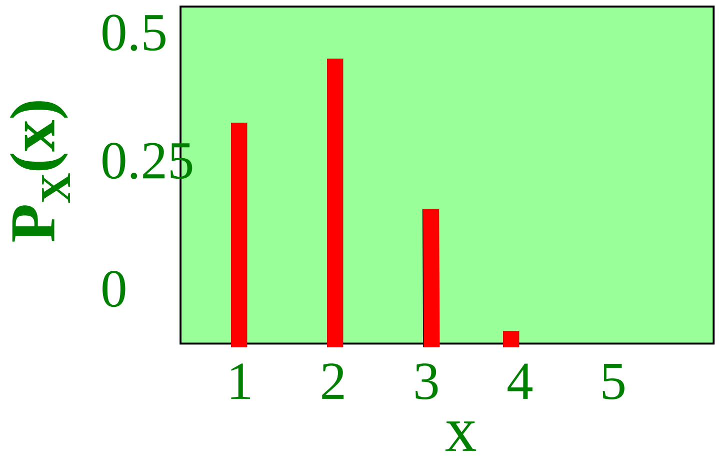
Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

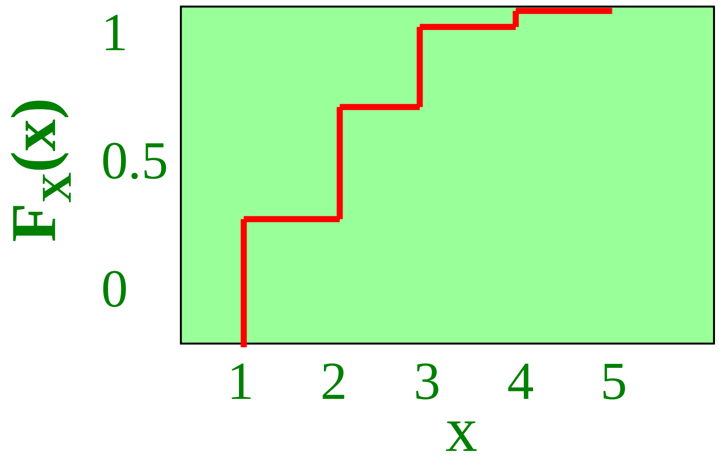
Mixed Random Variable

- Discrete RV → PMF & Summation
- Continuous RV → PDF & Integral
- Combination of Discrete and Continuous RV
  - →**Unit impulse function**
  - →**Can use same formulas to describe both RVs**

PMF → PDF

$$F_X(x) = \sum_{x_i \in S_X} P_X(x_i)\, u(x-x_i)$$



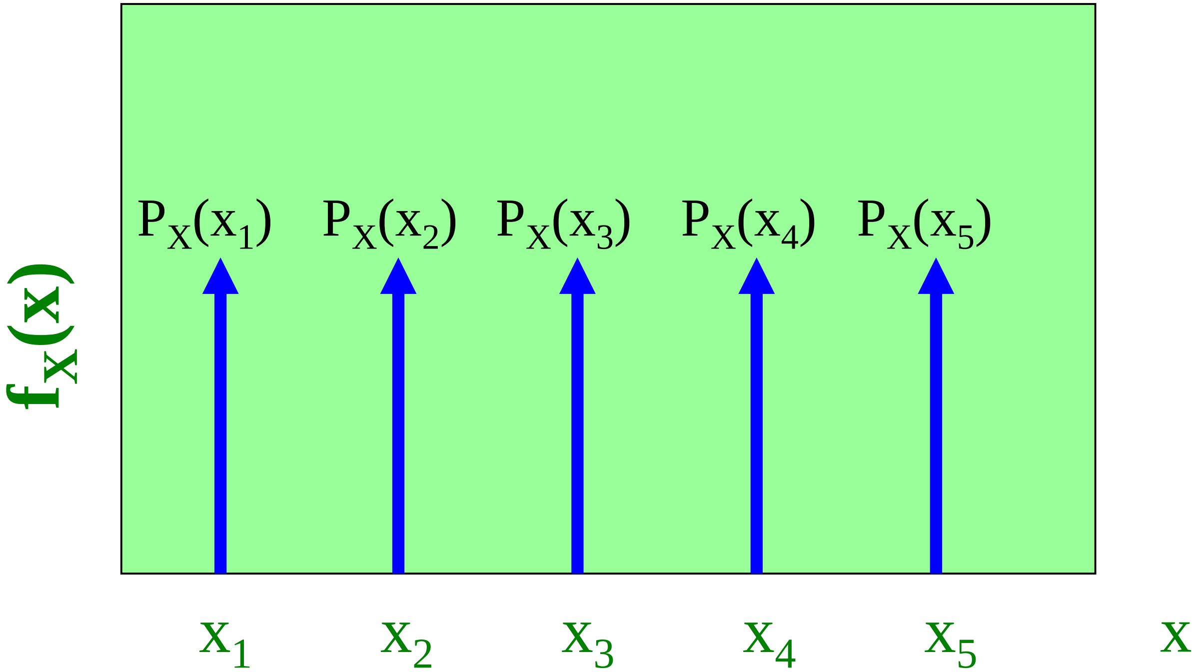


$u(x-x_i)$ → $u(x)$ shift to x_i

Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

PMF → PDF

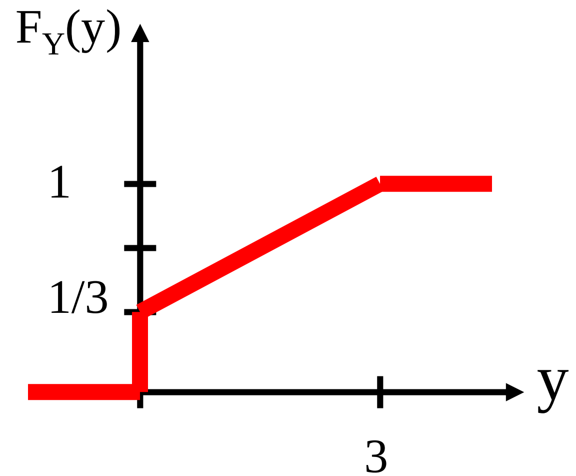
$$f_X(x) = \sum_{x_i \in S_X} P_X(x_i)\,\delta(x-x_i)$$



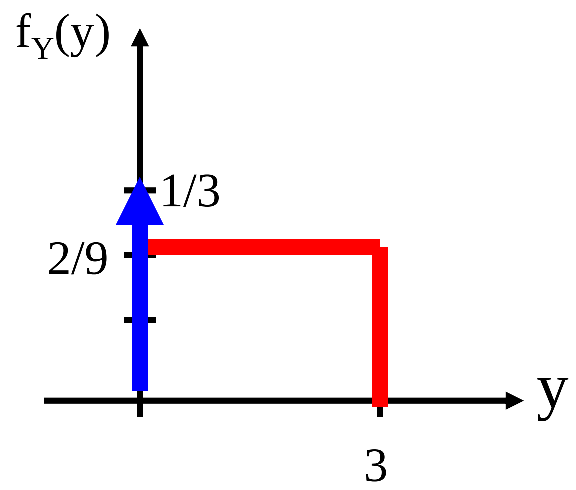
Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Example

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ 1/3 + 2y/9 & 0 \le y \le 3 \\ 0 & \text{Otherwise} \end{cases}$$



$$f_Y(y) = \begin{cases} \delta(y)/3 + 2/9 & 0 \le y \le 3 \\ 0 & \text{Otherwise} \end{cases}$$



Anan Phonphoem
Dept. of Computer Enginerring, Kasetsart University, Thailand

Summary

- Probability and Random Variable
- Discrete Random Variable
  - Uniform/Bernoulli/Geometric/…
  - PMF & CDF
  - Expected Value
  - Variance & Standard Deviation
- Continuous Random Variable
  - PDF
  - Uniform/Exponential/Gaussian

→Multiple Random Variables

→Stochastic Process

HW

- Summarize the "Memoryless Property" with your own word
- Give an clear example (with detail explanation) of the memoryless property
- Not more than 2 pages (A4)