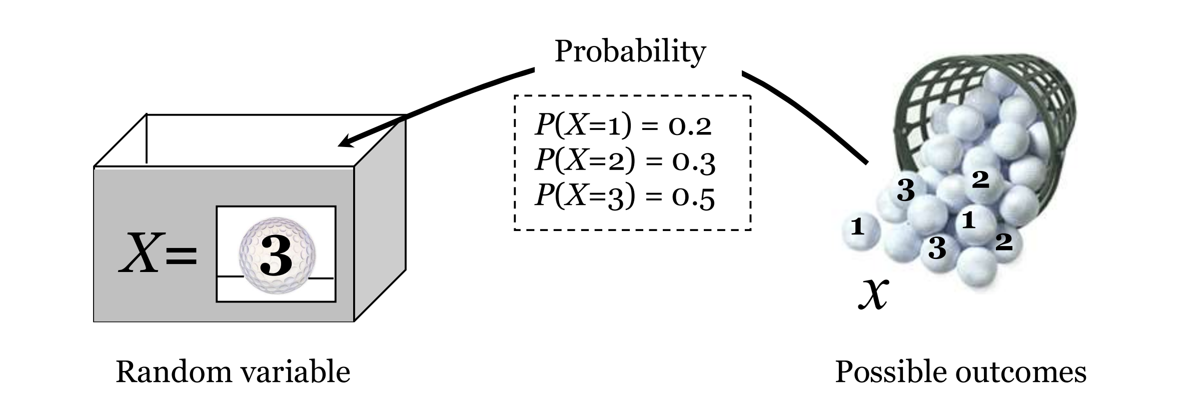
Seoul National University
Dept. of Civil and Environmental Engineering

Instructor: Junho Song
junhosong@snu.ac.kr

457.212 Statistics for Civil & Environmental Engineers
In-Class Material: Class 09
Random Variable and Probability Distribution Functions (A&T: 3.1)

1. **Random variable**

(a) Definition: a ~~variable~~ quantity that takes on any value in a specified set according to assigned probabilities.



(b) Discrete random variables

T: Number of tornadoes per year, {   }

B: Number of substandard beams out of 10, {   }

(c) Continuous random variables

C: Compact ratio of a soil specimen (0%: air only, 100%: no air), {c| }

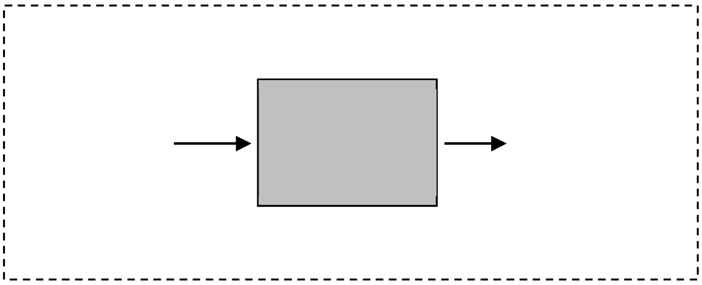
2. **Probability distribution functions**

	Discrete R.V.	Continuous R.V
Probability or Density at $X=x$	**P**robability **M**ass **F**unction (PMF), $P_X(x)$	**P**robability **D**ensity **F**unction (PDF), $f_X(x)$
Cumulative Frequency up to $X=x$	**C**umulative **D**istribution **F**unction (CDF), $F_X(x)$	

3. **Probability Mass Function** (PMF) of a **discrete** random variable X , $P_X(x)$

(a) Definition

$P_X(x) = P(\qquad\qquad\qquad)$



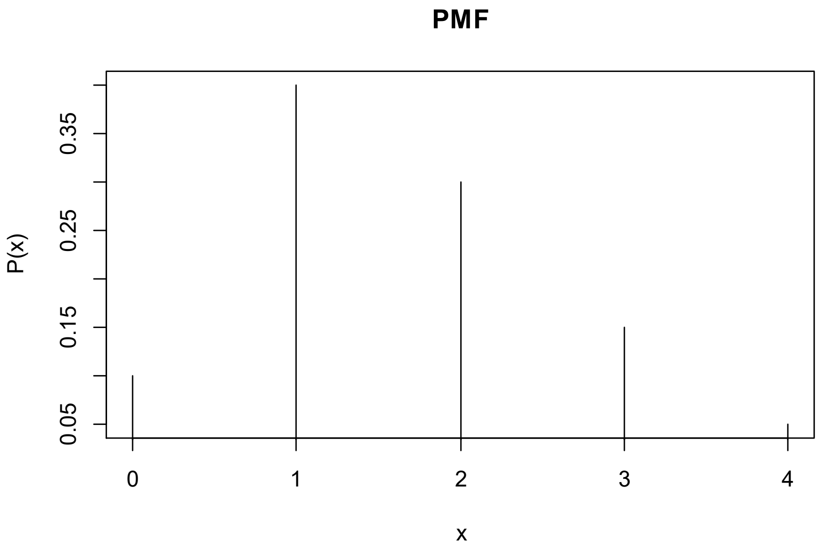
(b) Example

X : Number of landfalls of hurricanes per year

x	$P(X=x)=$
0	0.10
1	0.40
2	0.30
3	0.15
4	0.05

$P_X(2) =$

$P_X(4) =$



```
# Creating PMF plot by R
x = 0:4
Px = c(0.1,0.4,0.3,0.15,0.05)
plot(x,Px,type="h",xlab="x",ylab="P(x)",main="PMF")

# Use plot() with "h" type rather than barplot() to have a correct scale
on x-axis
x = c(0,1.5,2,3,4)
plot(x,Px,type="h",xlab="x",ylab="P(x)",main="PMF")
```

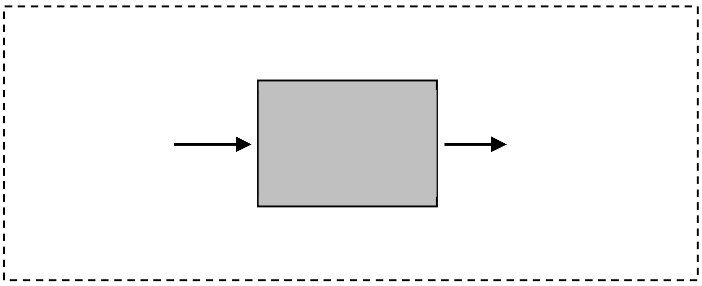
(c) Properties

- $\leq P_X(x) \leq$
- $\sum_{all\ x} P_X(x) =$, e.g. $P_X(0) + ... + P_X(4) =$
- $P(a < X \leq b) = \quad \sum P_X(x)$, e.g. $P(0 < X \leq 2) =$

4. **Cumulative Distribution Function** (CDF) of a **discrete** random variable X, $F_X(x)$

(a) Definition:

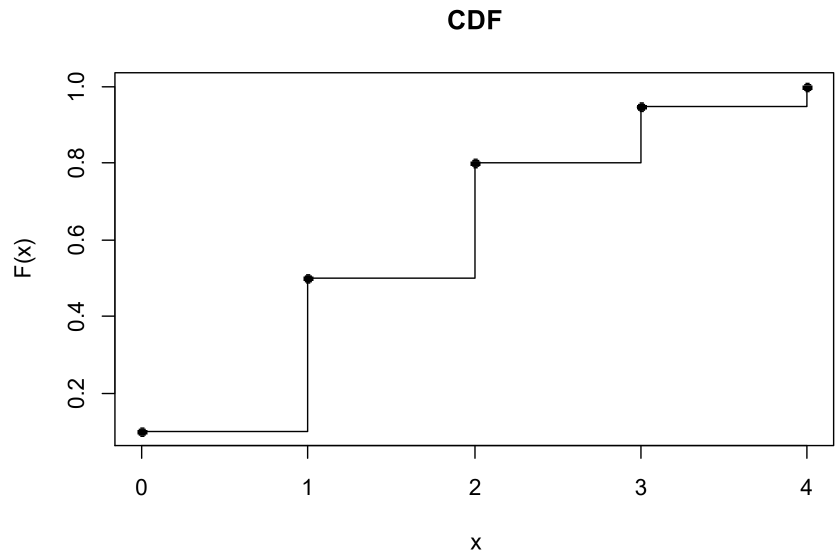
$F_X(x) = P(\qquad\qquad)$



(b) Example

X : Number of landfalls of hurricanes per year

x	$P_X(x)$	$F_X(x)=P(X\leq x)$
0	0.10	
1	0.40	
2	0.30	
3	0.15	
4	0.05	




```
# Creating CDF plot
Fx = cumsum(Px) # Creating CDF by cumulative sum of PMF
plot(x,Fx,type="s",xlab="x",ylab="F(x)",main="CDF")
lines(x,Fx,type="p",pch=16)
```

(c) Properties

- $F_X(a) = \sum P_X(x)$

- $F_X(-\infty) =$

  because $P(X \quad -\infty) =$

- $F_X(\infty) =$

  because $P(X \quad \infty) =$

- $P(a < x \le b) = \qquad\qquad -$

  e.g. $P(0 < X \le 2) =$

Instructor: Junho Song
junhosong@snu.ac.kr

Example 1: A company owns two buildings, A and B. The probabilities that the buildings A and B will be **non-operational** after an earthquake are 0.2 and 0.1, respectively. The buildings are located closely to each other. As a result, the probability that the building B is non-operational is twice the original probability if the building A is non-operational.

(a) What is the probability that there is no building in operation?

(b) What is the probability that both buildings are in operation?

(c) Let X denote the number of operational buildings. Find and plot the probability mass function (PMF) of X.

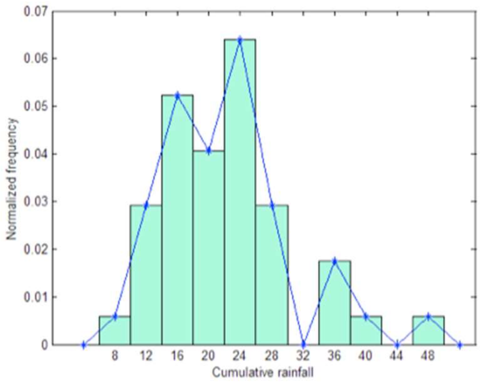
(d) Find and plot the cumulative distribution function (CDF) of X.

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

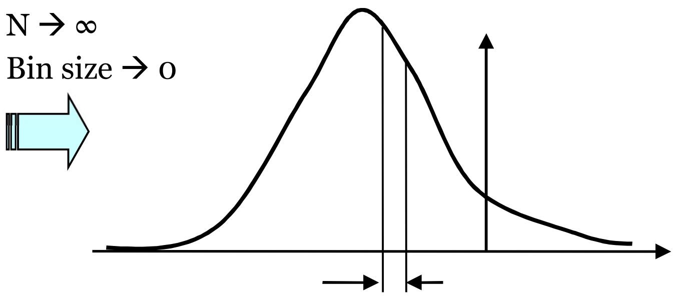
Instructor: Junho Song
junhosong@snu.ac.kr

5. **Probability Density Function** (PDF) of a **continuous** random variable X, $f_X(x)$

(a) Definition: "Density" of probability distribution at $X = x$.



Normalized Polygon



Probability Density Function

$$P(x < X \le x + \Delta x) = f_X(x) \cdot \Delta x$$

$$f_X(x) = \quad \lim \quad \underline{\hspace{6cm}}$$

(b) Properties

- $f_X(x) \qquad 0$
- $\int_{-\infty}^{\infty} f_X(x)dx =$; "Marginalization rule"
- $P(a < X \le b) = \int$