Probability and Random Process

Aimin Tang

The University of Michigan- Shanghai Jiao Tong University Joint Institute
Shanghai Jiao Tong University

Sep. 30 2021

Based on Lecture Notes by Prof. Chong Han

Outline

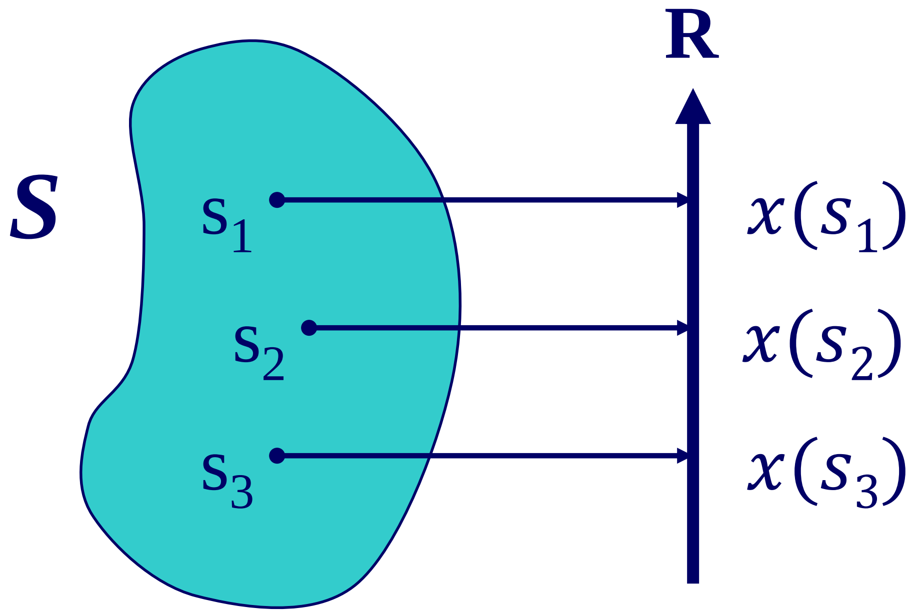
- 2. Random Variables
  - Introduction to Random Variables
  - PMF and Discrete Random Variables
  - PDF and Continuous Random Variables
  - Gaussian CDF
  - Conditional Probability
  - Function of a RV
  - Expectation of a RV
  - Transform Methods and Probability Generating Function

Introduction to Random Variables

Ve501 2021-2022 Fall

Random Variables

- A **random variable (RV)** is a **function** that maps outcomes in a sample space to the real numbers.

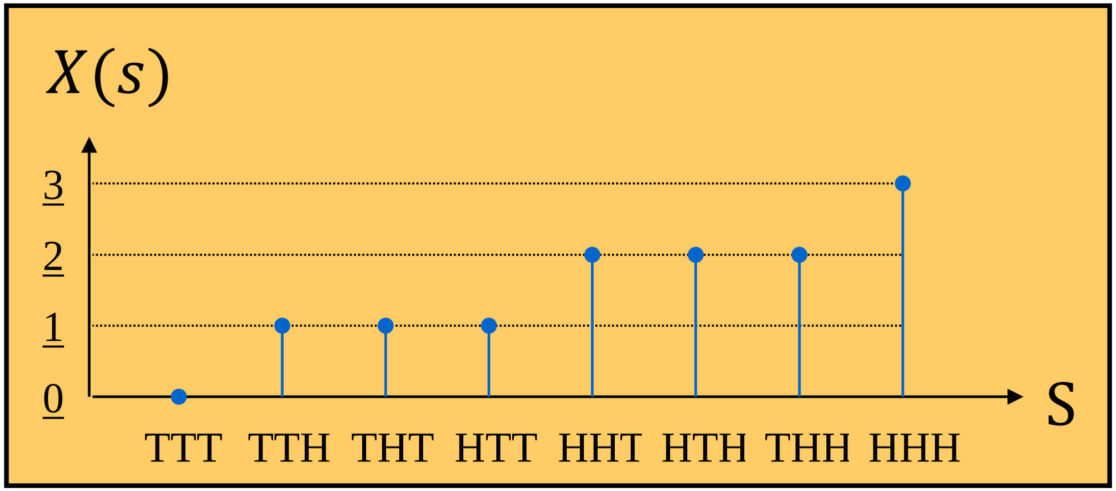


$$X(s) \in \mathbf{R}$$

- X has two meanings
  - X is a variable
  - X is a function

RV Example 1

- The sample space S comprises the ordered outcomes of tossing a fair coin three times.



Let $X(s)$ be the number of heads in three tosses.

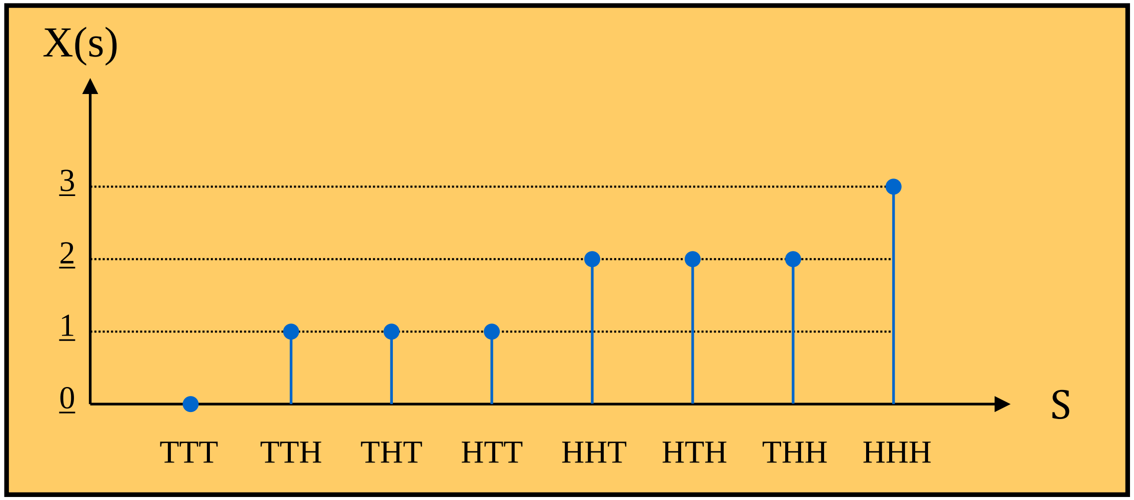
Conditions

- To be a random variable, a function must satisfy:

1. The event $\{X(s) \leq x\}$ must correspond to a valid event on S (i.e. a member of the field of events in the probability triplet) for every $x \in \mathbf{R}$.

2. $\Pr(X(s) = +\infty) = \Pr(X(s) = -\infty) = 0$

Return to Example 1



Let the event B be $B = \{X(s) \le 1.5\}$

$$B = \{X(s) \le 1.5\} = \{TTT, TTH, THT, HTT\}$$

Then,

$$P(B) = \frac{1}{2}$$

Cumulative Distribution Function

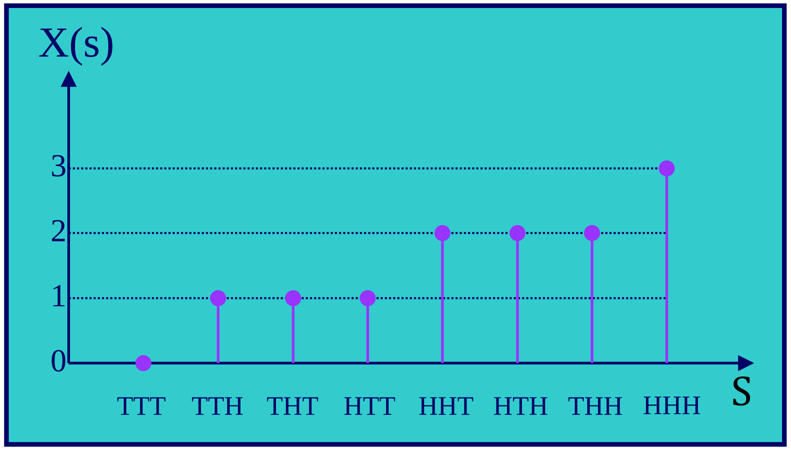
- The cumulative distribution function (CDF) is a real-valued function on $\boldsymbol{R}$, denoted $F_X(x)$, and defined

$$F_X(x) = \Pr(X(s) \leq x)$$

Subscript names the random variable

x is just a "dummy variable" that is used as a threshold

Evaluating the CDF



$$F_X(3) =$$

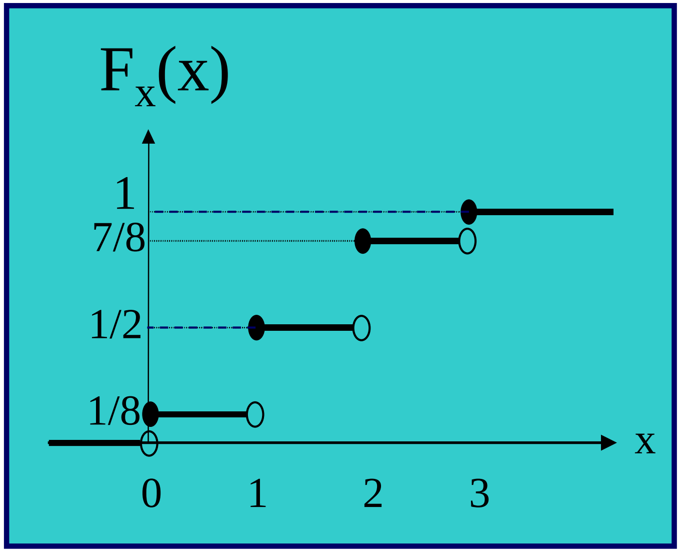
$$F_X(2.9)$$

$$F_X(1.9) =$$

$$F_X(0.9) =$$

$$F_X(-0.1)$$

The CDF from Example 1



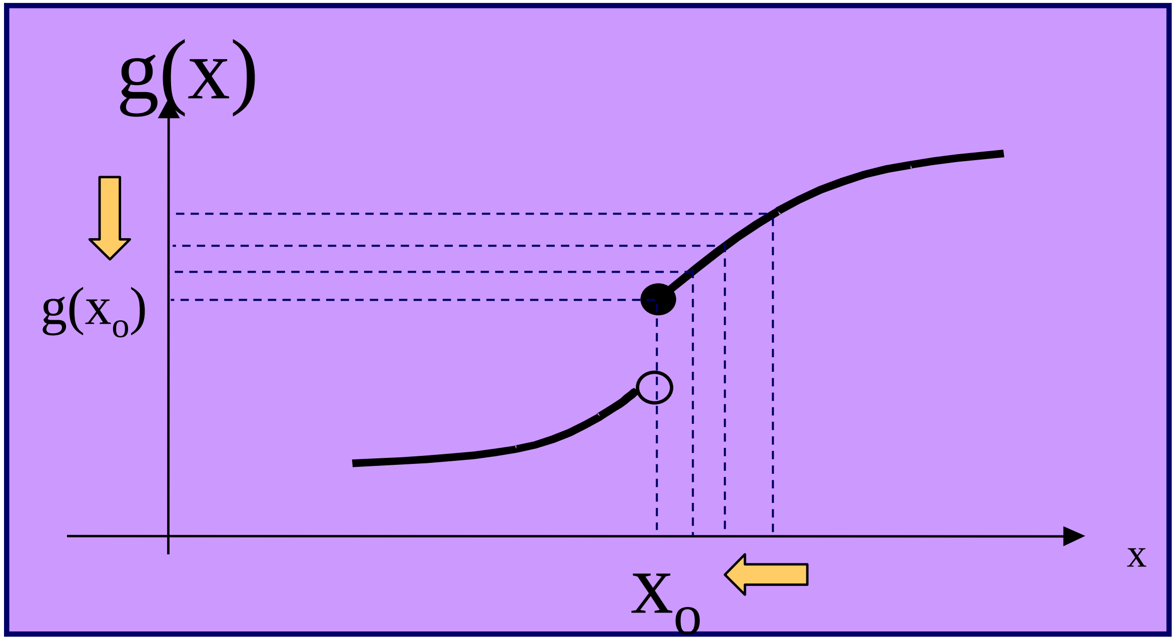
$$F_X(x) = \begin{cases} 0, & x < 0 \\ 1/8, & 0 \le x < 1 \\ 1/2, & 1 \le x < 2 \\ 7/8, & 2 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$

Observe that $F_X(x)$ is **continuous from the right**.

Right Continuity

- A function g(x) is continuous from the right at x_0 when

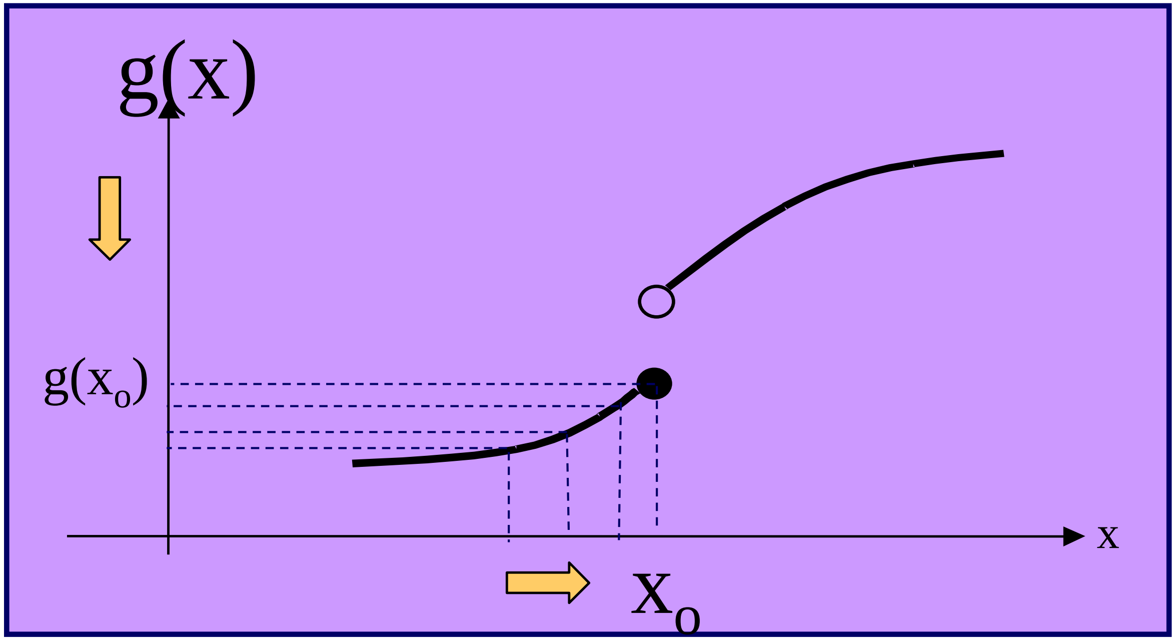
$$\lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} g(x_0 + \varepsilon) = g(x_0)$$



Left Continuity

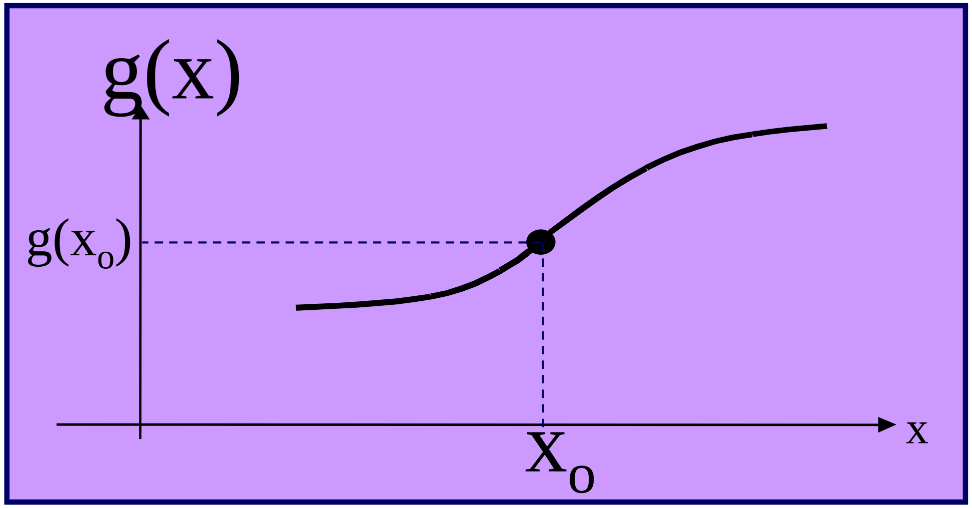
- $g(x)$ is continuous from the left at x_0 when

$$\lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} g(x_0 - \varepsilon) = g(x_0)$$



Continuity

- $g(x)$ is continuous at x_0 if it is both left-continuous and right continuous at x_0.



- $g(x)$ is a continuous function if it is continuous for all $x \in \boldsymbol{R}$.

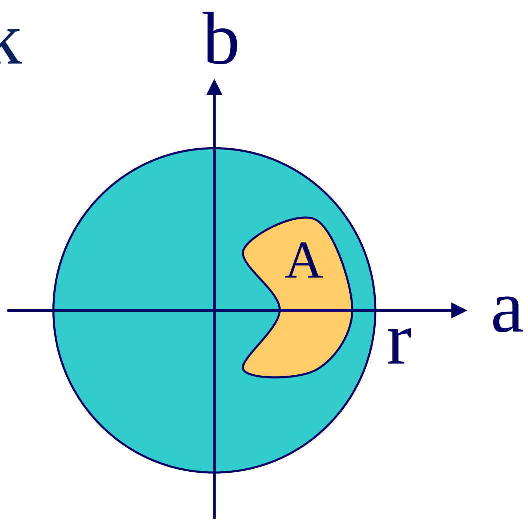
Random Variable Example -- I

- A point is chosen at random on a disk of radius r:

$$S = \{(a,b): \sqrt{a^2+b^2} \le r\}$$

- "At random" means that if A is some subset of Ω, then

$$\Pr\big((a,b) \in A\big) = \frac{\text{area of } A}{\pi r^2}$$

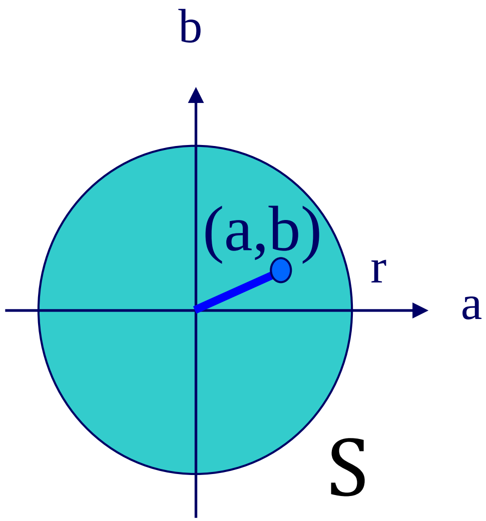


Random Variable Example -- II

- Let the random variable X be defined

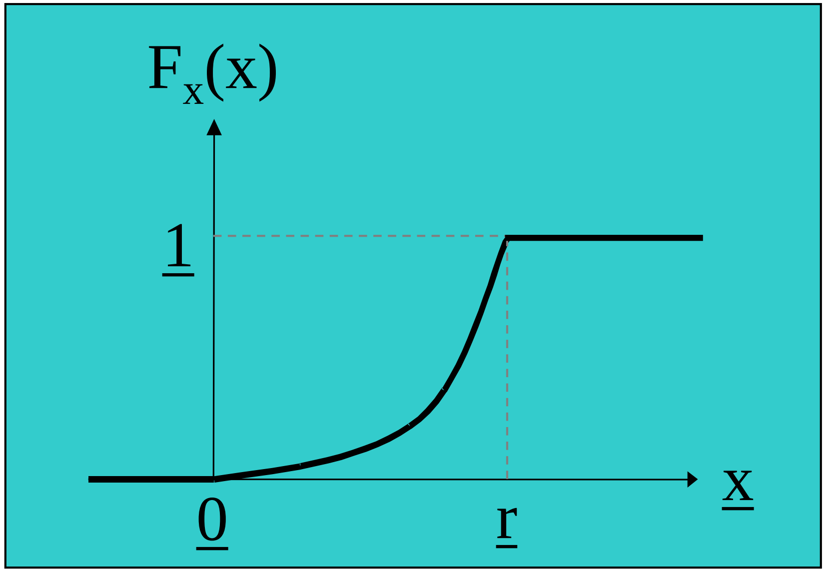
$$X(s) = \sqrt{a^2 + b^2}$$

= distance of point s with coordinates (a,b), from the origin



CDF for Ex. 2

$$F_X(x) = P[X(s) \le x]$$



Properties of CDFs

1. $0 \le F_X(x) \le 1$ for all $x \in R$

2. $\lim\limits_{x \to -\infty} F_X(x) = 0, \quad \lim\limits_{x \to +\infty} F_X(x) = 1$

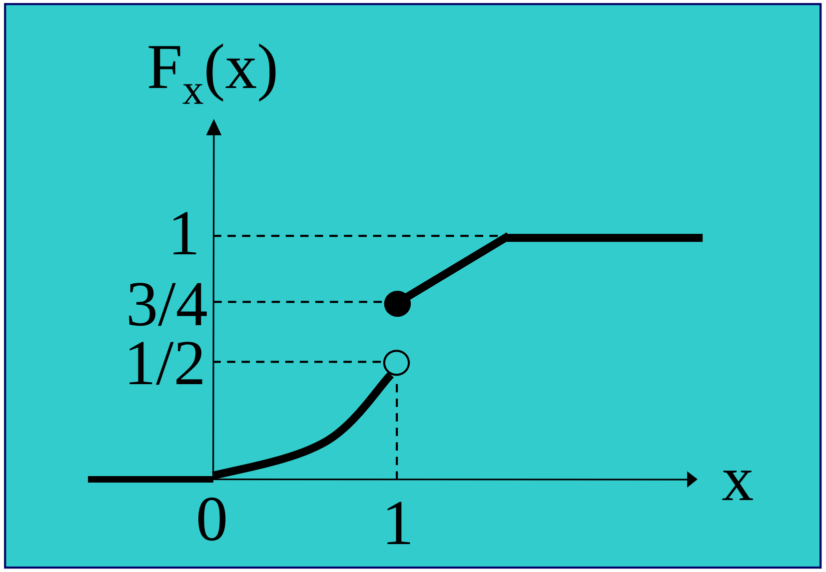
3. $F_X(x)$ is non-decreasing.

4. $F_X(x)$ is right-continuous.

5. $P(X(s) < x_0) = \lim\limits_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} F_X(x_0 - \varepsilon) =$ limit from the left at x_0.

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

Example of CDF Property 5



$$P(X \leq 1) = F_X(1) = \frac{3}{4}$$

$$P(X < 1) = \lim_{\substack{\epsilon > 0 \\ \epsilon \to 0}} F_X(1 - \epsilon) = \frac{1}{2}$$

Proof of CDF Property 6

Let a < b, then

$$\{X(s) \leq b\} = \{X(s) \leq a\} \cup \{a < X(s) \leq b\}$$

Disjoint events

It follows that

$$\Pr(X(s) \leq b) = \Pr(X(s) \leq a) + \Pr(a < X(s) \leq b)$$

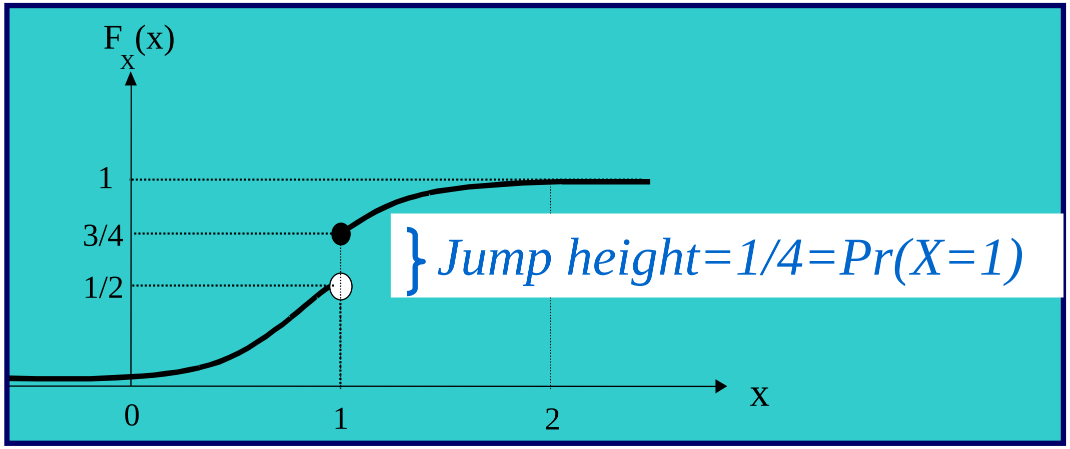
Using the definition of $F_X(x)$,

$$F_X(b) = F_X(a) + \Pr(a < X(s) \leq b)$$

Equality

$$P(X=a) = \text{ discontinuity height at } a \;=\; F_X(a) - \lim_{\substack{\varepsilon>0 \\ \varepsilon\to 0}} F_X(a-\varepsilon)$$

$$P(X=1) \;=\; 3/4 - 1/2 \;=\; 1/4$$



- Simplification for notation: there is no s-dependence indicated by the notation - it is still assumed. X is just simpler to write than $X(s)$

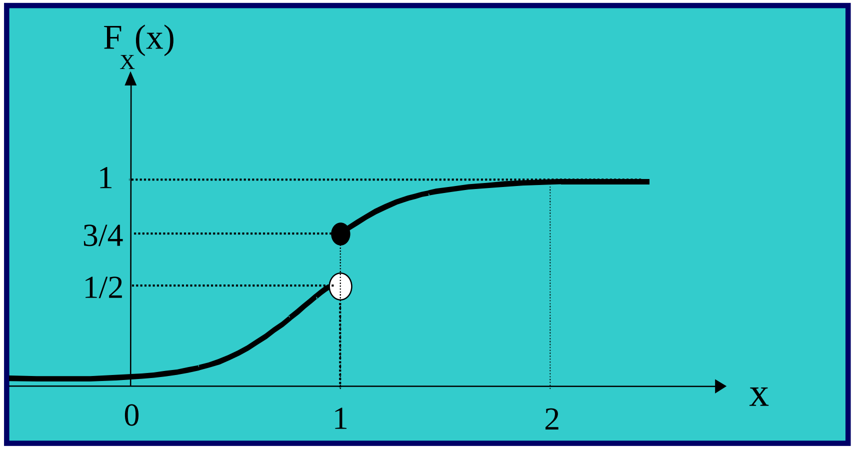
Proof of $\mathrm{P}(X = a)$

$$\{X \le a\} = \{X < a\} \cup \{X = a\}$$

$$P(X \le a) = P(X < a) + P(X = a)$$

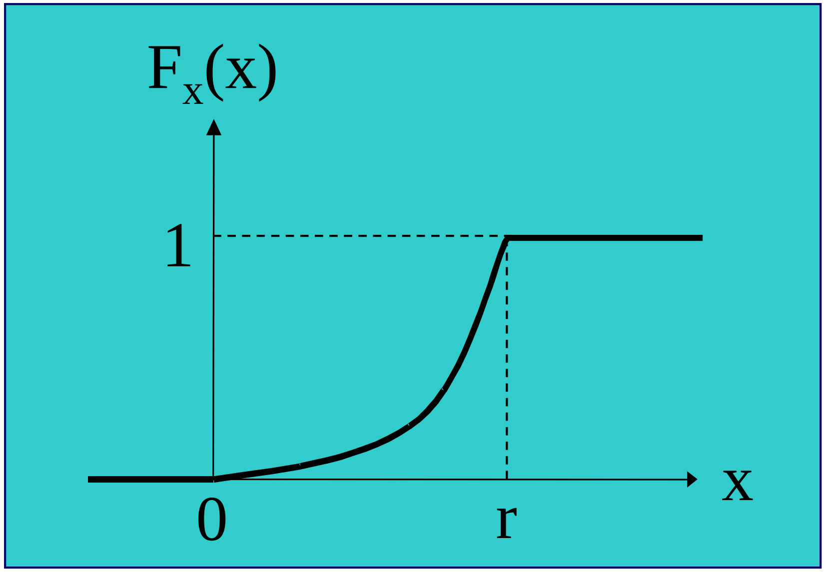
$$F_X(a) = \lim_{\substack{\varepsilon > 0 \\ \varepsilon \to 0}} F_X(a - \varepsilon) + P(X = a)$$

Therefore, $P(X = a) = F_X(a) - \lim_{\substack{\varepsilon > 0 \\ \varepsilon \to 0}} F_X(a - \varepsilon)$



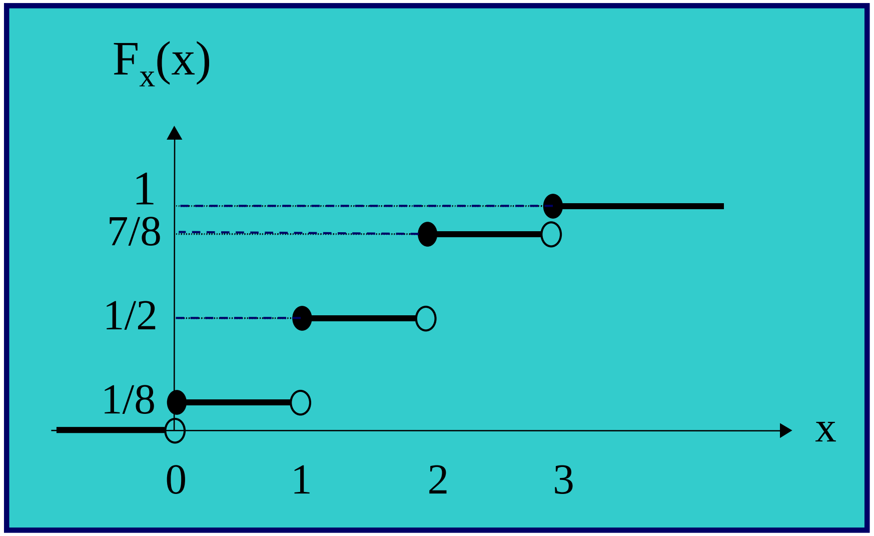
Continuous Random Variables

- If $F_X(x)$ is a continuous function (i.e. continuous at all $x \in \boldsymbol{R}$), then X is a **continuous random variable**



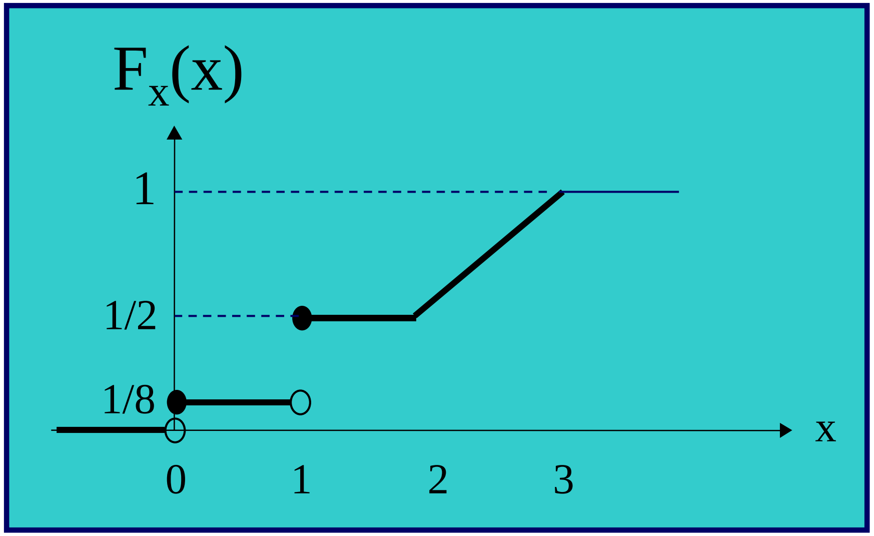
Discrete Random Variables

- If $F_X(x)$ is a piecewise constant function (i.e. flat everywhere except at discontinuities), then X is a **discrete random variable**



Mixed Random Variables

- If X is neither continuous nor discrete, then it is mixed



Short Summary

- A random variable (RV) is a function
- The RV is described by its cumulative distribution function (CDF)
- Continuity
  - CDFs are right-continuous
- CDFs properties
- Three classes of RVs:
  - Continuous
  - Discrete
  - Mixed

PMF and Discrete Random Variables

The Probability Mass Function

- An alternative description of a **discrete random variable** is the **probability mass function (PMF)**.
- The discrete RV maps outcomes in a sample space to the real numbers.
- The PMF indicates the probability that the random variable exactly equals some values:

$$p_X(x) = \Pr(X = x)$$

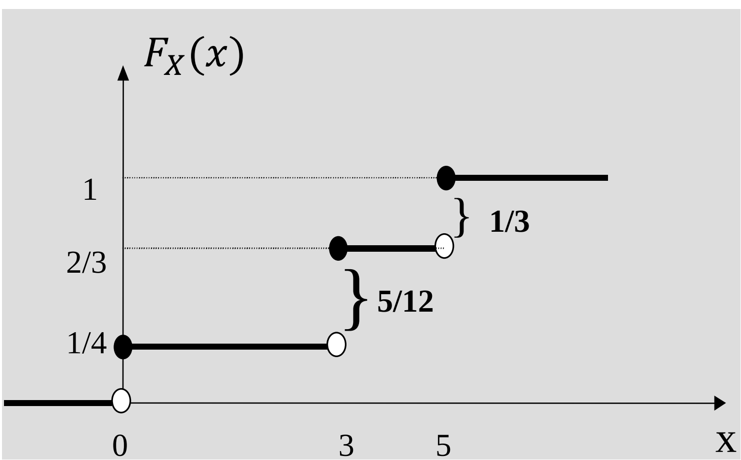
- The plot of $p_X(x)$ is a stem plot, where the sum of the stems is 1.

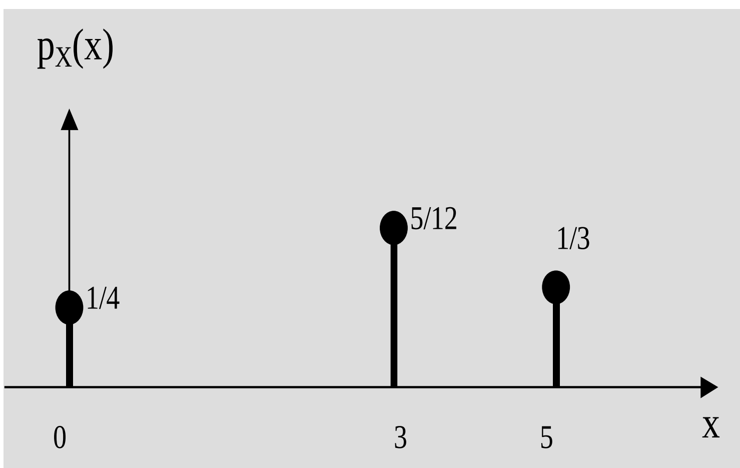
PMF Example

- Recall that: $$Pr(X = x_0) = F_X(x_0) - \lim_{\substack{\varepsilon>0 \\ \varepsilon\to 0}} F_X\,(x_0 - \varepsilon)$$

- In other words,

$$\Pr(X = x_0) = \text{the height of the jump in the CDF at } x = x_0$$





Some Special Discrete Random Variables

- Uniform
- Bernoulli
- Binomial
- Geometric
- Poisson

The Uniform Random Variable

- $X(s)$ is the Uniform RV if $A = \{1,2, \dots n\}$ and

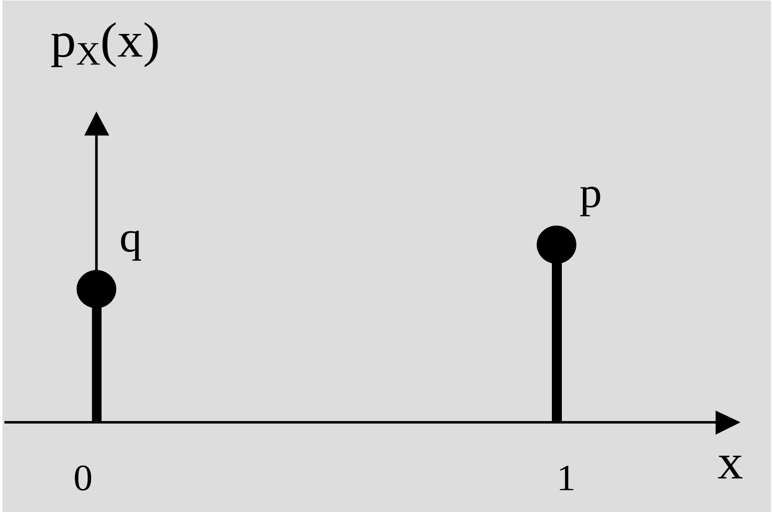
$$p_X(k) = \frac{1}{n}, k = 1,2, \dots n$$

- Example
  - random number generator
  - toss a fair dice
  - draw a card from a well-shuffled deck of cards

The Bernoulli Random Variable

- Let A be an event on S. $X(s)$ is the Bernoulli RV that indicates A if:

$$X(s) = \begin{cases} 1 \text{ if } s \in A \\ 0 \text{ if } s \notin A \end{cases}$$



$$p_X(0) = Pr(X = 0) = q$$
$$p_X(1) = Pr(X = 1) = p$$

The Binomial Random Variable

- Consider n Bernoulli Trials and let S be the Cartesian product sample space for all n trials.

  X(s) number of successes in s where $s \in S$

  $$p_X(k) = \Pr(k \text{ successes in } n \text{ trails})$$
  $$= \binom{n}{k} p^k q^{n-k}$$

  Where the probability of success in one try in p

  - The binomial distribution with parameters n and p is the discrete probability distribution of the number of successes in a sequence of n independent yes/no experiments, each of which yields success with probability p.
  - When n=1, Binomial is equivalent to Bernoulli

Binomial Properties

$p_X(k)$ takes its maximum value at $k_{\max} = \lfloor (n+1)p \rfloor$,

where the floor function $\lfloor x \rfloor$ = the greatest integer $\leq x$

- When $(n+1)p$ is an integer, the maximum value is achieved at both $k_{\max}$ and $k_{\max}$-1. Can you prove this?

The Geometric Random Variable

- Consider an infinite sequence of Bernoulli Trials, $X(s)$ is the number of failures before the first success
  - $p_X(k) = q^k p, k = 0,1,2, \ldots$
  - Where the probability of success in one try in p
- Question: can you calculate the summation of $p_X(k)$?
  - $\sum_{k=0}^{\infty} p_X(k) = (1-q)\sum_0^{\infty} q^k = (1-q)\frac{1}{1-q} = 1$
- Application:
  - In a memoryless binary communication link, where q is bit error, $p_X(k)$describes the probability of getting a burst of errors k bits long.

The Poisson Random Variable

- In an application where events happen at random points in time, this RV counts the number of events that occur in a specified time interval.
  - Packet arrivals in a computer network
  - Customer arrivals
  - Lighting strikes
  - Photon arrivals
  - Component failures

Poisson PMF

- X is a **Poisson random variable** with parameter $\Lambda > 0$ if $A = \{0, 1, 2, \dots\}$ and

$$p_X(k) = \begin{cases} \dfrac{e^{-\Lambda}\Lambda^k}{k!}, & k \in A \\ 0, & \text{otherwise} \end{cases}$$

- Λ is the "average" number of occurrences in the time interval T
- Λ can be expressed as λT
  - where λ is the "average rate" of occurrences

Remarks

- Note:

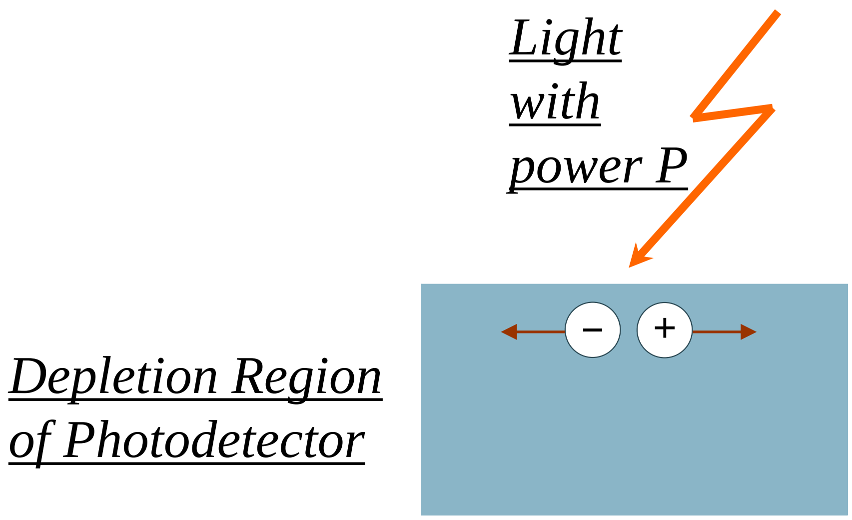
$$\sum_{k=0}^{\infty} \frac{\Lambda^k}{k!} = e^{\Lambda}$$

- Thus,

$$\sum_{k=0}^{\infty} p_X(k) = \sum_{k=0}^{\infty} \frac{e^{-\Lambda}\Lambda^k}{k!} = e^{-\Lambda}e^{\Lambda} = 1$$

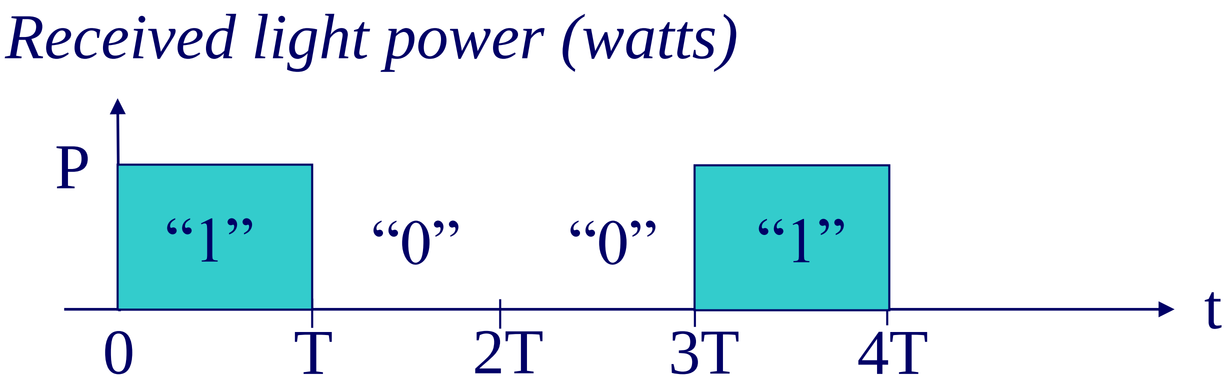
Example Poisson Application: Photodetection

- When a photon of light energy falls on a photodetector, its energy is either absorbed by the lattice or it creates an electron-hole (E-H) pair.
- Because a photodetector is a reverse-biased diode, the E-H pair immediately separates in the depletion region, creating a small current.



Photon-Counting Receiver Operation

- An idealized receiver used as a “benchmark” for performance.



- In the receiver, a “one” is declared if at least one E-H pair is created in the interval *T*. A “zero” is declared otherwise.

Poisson Approximation to Binomial

- Divide time line into n equal intervals Δ wide. Δ is small enough that $\lambda\Delta$ <<1.

- Success = at least one event occurs in an interval $n\Delta$

- Good when n is large, p is small, and np is on the order of $n\lambda\Delta$ infrequent success

Poisson Approximation to Binomial

- Success is when at least one event occurs in an interval $n\Delta$

$$p = 1 - e^{-\lambda\Delta} \approx \lambda\Delta$$

$$q \approx 1 - \lambda\Delta$$

$$P(\text{k events in } n\Delta) = \frac{(\lambda n\Delta)^k}{k!} e^{-\lambda n\Delta}$$

$$\approx \binom{n}{k} (\lambda\Delta)^k (1 - \lambda\Delta)^{n-k}$$

Accuracy of Approximation

- N=20, p=0.08

Binomial Distribution* ($n = 20 \quad p = .08$)	Poisson Distribution† ($\lambda = np = 1.6$)
$P(X=0) = .1887$	$P(X=0) = \frac{e^{-1.6}(1.6)^0}{0!} = .2019$
$P(X=1) = .3282$	$P(X=1) = \frac{e^{-1.6}(1.6)^1}{1!} = .3230$
$P(X=2) = .2711$	$P(X=2) = \frac{e^{-1.6}(1.6)^2}{2!} = .2584$
$P(X=3) = .1414$	$P(X=3) = \frac{e^{-1.6}(1.6)^3}{3!} = .1378$
$P(X=4) = .0523$	$P(X=4) = \frac{e^{-1.6}(1.6)^4}{4!} = .0551$
$P(X=5) = .0145$	$P(X=5) = \frac{e^{-1.6}(1.6)^5}{5!} = .0176$
$P(X=6) = .0032$	$P(X=6) = \frac{e^{-1.6}(1.6)^6}{6!} = .0047$
$P(X=7) = .0005$	$P(X=7) = \frac{e^{-1.6}(1.6)^7}{7!} = .0011$
$P(X=8) = .0001$	$P(X=8) = \frac{e^{-1.6}(1.6)^8}{8!} = .0002$
$P(X=9) = .0000$	$P(X=9) = \frac{e^{-1.6}(1.6)^9}{9!} = .0000$
$P(X=10) = .0000$	$P(X=10) = \frac{e^{-1.6}(1.6)^{10}}{10!} = .0000$
$P(X=20) = \underline{.0000}$	$P(X=20) = \frac{e^{-1.6}(1.6)^{20}}{20!} = \underline{.0000}$

Short Summary

- The discrete RVs covered were:
  - Uniform
  - Bernoulli
  - Binomial
  - Geometric
  - Poisson
- The Poisson PMF can be used to approximate the Binomial PMF

Quiz

- The number of hits to a popular website during a 5-minute interval is given by a Poisson random variable with $\Lambda = 1$. Find the probability that there are at least 3 hits between 3:00am and 3:05am.

Solution

- The number of hits to a popular website during a 5-minute interval is given by a Poisson random variable with $\Lambda = 1$. Find the probability that there are at least 3 hits between 3:00am and 3:05am.

- Solution:

$$\Pr(X \geq 3) = \sum_{k=3}^{\infty} \frac{\Lambda^k e^{-\Lambda}}{k!}$$

$$= 1 - \Pr(X = 0) - \Pr(X = 1) - \Pr(X = 2)$$

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

$$= 1 - e^{-1}\left(1 + 1 + \frac{1}{2}\right) \approx 0.08$$

Thank You!