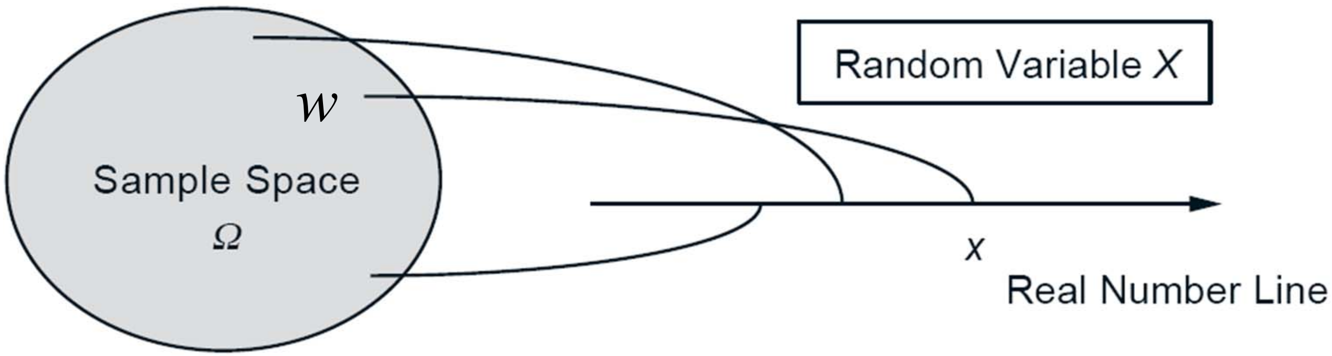
Discrete Random Variables: Basics

Berlin Chen
Department of Computer Science & Information Engineering
National Taiwan Normal University

Reference:
- D. P. Bertsekas, J. N. Tsitsiklis, *Introduction to Probability* , Sections 2.1-2.3

Random Variables

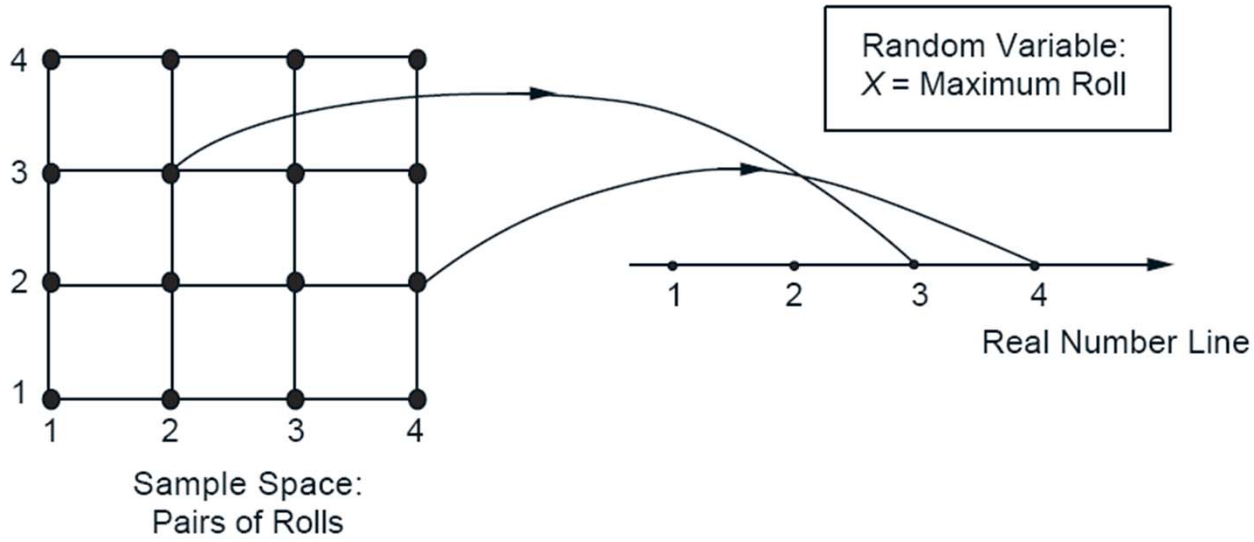
- Given an experiment and the corresponding set of possible outcomes (the sample space), a random variable associates a particular number with each outcome
  - This number is referred to as the (numerical) value of the random variable
  - We can say a random variable is a real-valued function of the experimental outcome



$$X : w \rightarrow x$$

Random Variables: Example

- An experiment consists of two rolls of a 4-sided die, and the random variable is the maximum of the two rolls
  - If the outcome of the experiment is (4, 2), the value of this random variable is 4
  - If the outcome of the experiment is (3, 3), the value of this random variable is 3



  - Can be one-to-one or many-to-one mapping

Main Concepts Related to Random Variables

- For a probabilistic model of an experiment
  - A random variable is **a real-valued function** of the outcome of the experiment

$$X : w \rightarrow x$$

  - A **function of a random variable** defines another random variable

$$Y = g(X)$$

  - We can associate with each random variable certain "averages" of interest such the **mean** and the **variance**
  - A random variable can be **conditioned** on an event or on another random variable
  - There is a notion of **independence** of a random variable from an event or from another random variable

Discrete/Continuous Random Variables

- A random variable is called **discrete** if its **range** (the set of values that it can take) is finite or at most countably infinite

$$\text{finite}: \{1, 2, 3, 4\}, \text{countably infinite}: \{1, 2, \cdots\}$$

- A random variable is called **continuous (not discrete)** if its **range** (the set of values that it can take) is uncountably infinite
  - E.g., the experiment of choosing a point a from the interval $[-1, 1]$
    - A random variable that associates the numerical value a^2 to the outcome a is not discrete

- In this chapter, we focus exclusively on discrete random variables

Concepts Related to Discrete Random Variables

- For a probabilistic model of an experiment
  - A **discrete random variable** is a real-valued function of the outcome of the experiment that can take a finite or countably infinite number of values
  - A (discrete) random variable has an associated **probability mass function** (PMF), which gives the probability of each numerical value that the random variable can take
  - A **function of a random variable** defines another random variable, whose PMF can be obtained from the PMF of the original random variable

Probability Mass Functions

- A (discrete) random variable X is characterized through the probabilities of the values that it can take, which is captured by the probability mass function (PMF) of X, denoted $p_X(x)$

$$p_X(x) = \mathbf{P}(\{X = x\}) \text{ or } p_X(x) = \mathbf{P}(X = x)$$

  - The sum of probabilities of all outcomes that give rise to a value of X equal to x
  - **Upper case** characters (e.g., X) denote random variables, while **lower case** ones (e.g., x) denote the numerical values of a random variable

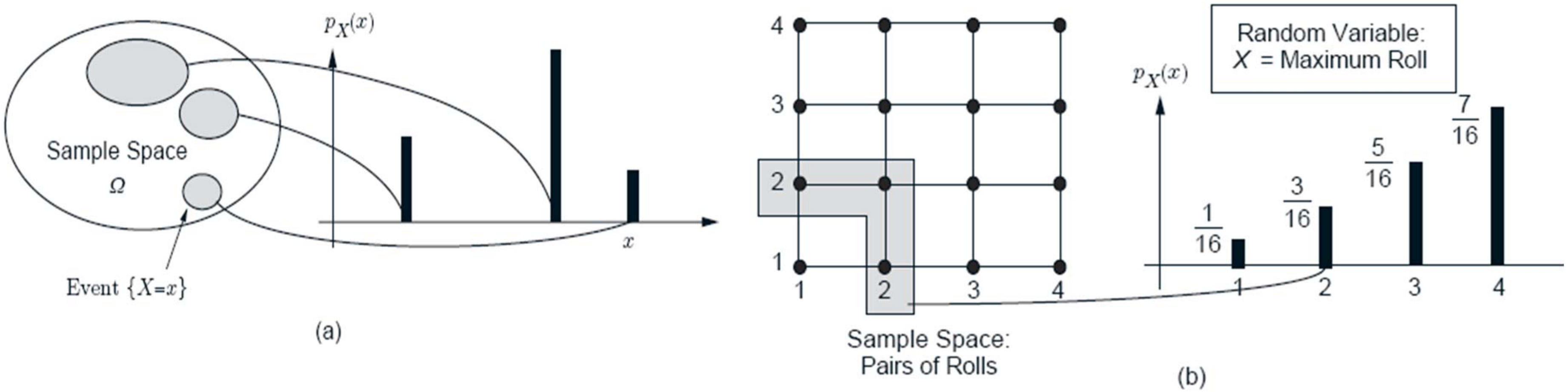
- The summation of the outputs of the PMF function of a random variable over all it possible numerical values is equal to one $\sum_x p_X(x) = 1$

$\{X = x\}$'s are disjoint and form a partition of the sample space

Calculation of the PMF

- For each possible value x of a random variable X :
  1. Collect all the possible outcomes that give rise to the event $\{X = x\}$
  2. Add their probabilities to obtain $p_X(x)$

- An example: the PMF $p_X(x)$ of the random variable X = maximum roll in two independent rolls of a fair 4-sided die



Bernoulli Random Variable

- A Bernoulli random variable X takes two values 1 and 0 with probabilities p and $1-p$, respectively
  - **PMF**

$$p_X(x) = \begin{cases} p, & \text{if } x = 1 \\ 1-p, & \text{if } x = 0 \end{cases}$$

- The Bernoulli random variable is often used to model generic probabilistic situations with just two outcomes

    1. The toss of a coin (outcomes: head and tail)
    2. A trial (outcomes: success and failure)
    3. the state of a telephone (outcomes: free and busy)

    …

Binomial Random Variable (1/2)

- A binomial random variable X has parameters n and p
  - PMF

$$p_X(k) = \mathbf{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0,1,\ldots,n$$

- The Bernoulli random variable can be used to model, e.g.
  1. The number of heads in *n* independent tosses of a coin (outcomes: 1, 2, …,*n*), each toss has probability p to be a head
  2. The number of successes in *n* independent trials (outcomes: 1, 2, …,*n*), each trial has probability p to be successful

- Normalization Property

Note that : $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$

$$\sum_{k=0}^{n} p_X(k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = 1$$

Binomial Random Variable (2/2)

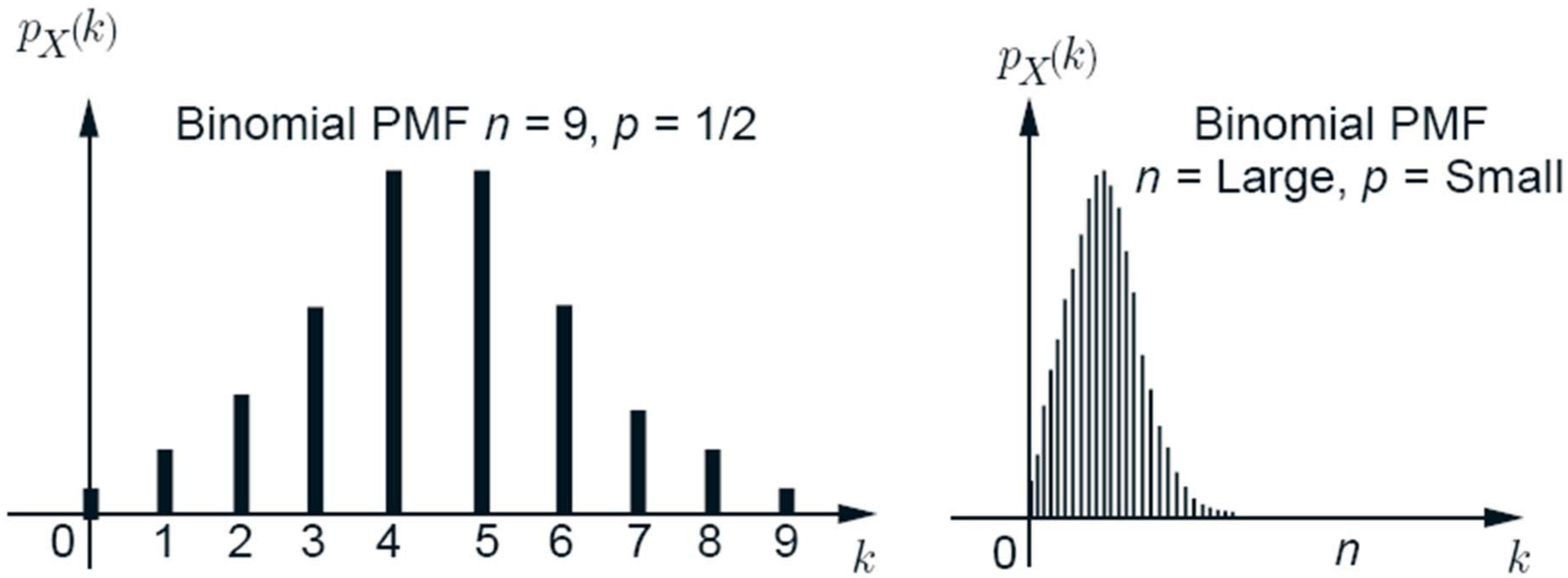
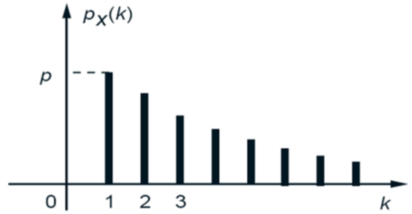


Figure 2.3: The PMF of a binomial random variable. If $p = 1/2$, the PMF is symmetric around $n/2$. Otherwise, the PMF is skewed towards 0 if $p < 1/2$, and towards n if $p > 1/2$.

Geometric Random Variable

- A geometric random variable X has parameter p $(0 < p < 1)$
  - PMF

$$p_X(k) = (1-p)^{k-1} p, \quad k = 1,2,\ldots,$$



- The geometric random variable can be used to model, e.g.
  - The number of independent tosses of a coin needed for a head to come up for the first time, each toss has probability p to be a head
  - The number of independent trials until (and including) the first "success", each trial has probability p to be successful
- Normalization Property

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

Poisson Random Variable (1/2)

- A Poisson random variable X has parameter λ
  - PMF

$$p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k = 0,1,2,\ldots,$$

- The Poisson random variable can be used to model, e.g.
  - The number of typos in a book
  - The numbers of cars involved in an accidents in a city on a given day

- Normalization Property

McLaurin series

$$\sum_{k=0}^{\infty} p_X(k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \Bigg(\underbrace{1 + \lambda + \frac{\lambda^2}{2!} + \frac{\lambda^3}{3!} + \cdots}_{e^{\lambda}} \Bigg) = 1$$

Poisson Random Variable (2/2)

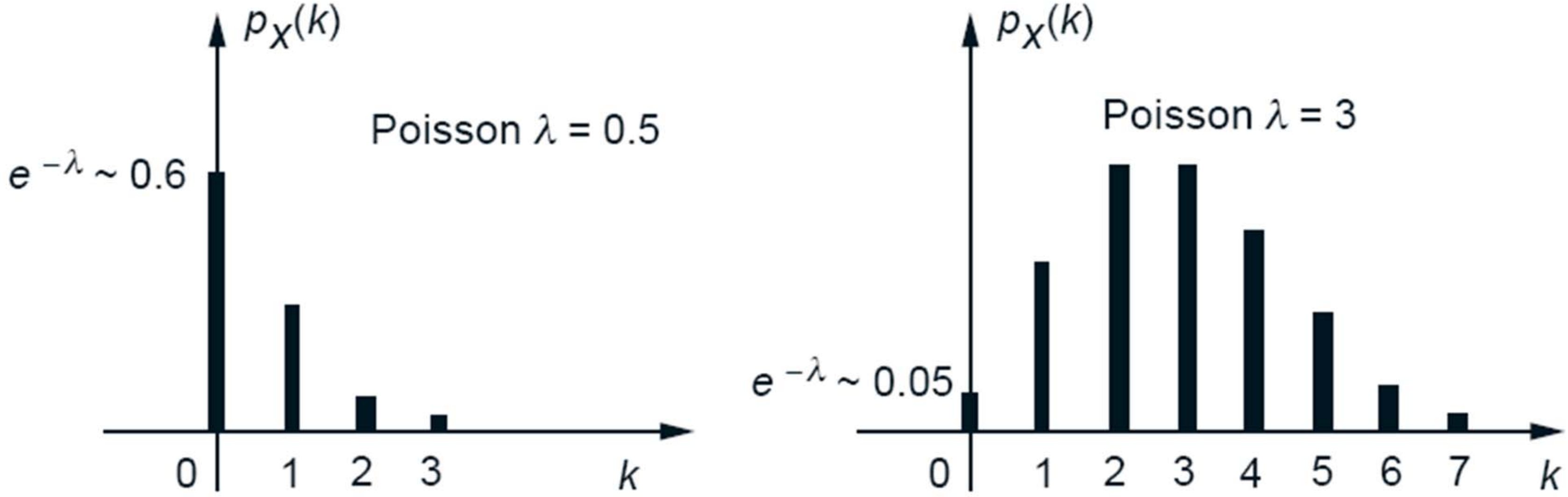


Figure 2.5: The PMF $e^{-\lambda}\frac{\lambda^k}{k!}$ of the Poisson random variable for different values of λ. Note that if $\lambda < 1$, then the PMF is monotonically decreasing, while if $\lambda > 1$, the PMF first increases and then decreases as the value of k increases (this is shown in the end-of-chapter problems).

Relationship between **Binomial** and **Poisson**

- The Poisson PMF with parameter λ is a good approximation for a binomial PMF with parameters n and p, provided that $\lambda = np$, n is very large and p is very small

$$\lim_{n\to\infty}\binom{n}{k}p^k(1-p)^{n-k}$$

$$= \lim_{n\to\infty}\frac{n!}{(n-k)!k!}p^k(1-p)^{n-k} \qquad (\because \lambda = np \Rightarrow p = \frac{\lambda}{n})$$

$$= \lim_{n\to\infty}\frac{n(n-1)\cdots(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \lim_{n\to\infty}\frac{\lambda^k}{k!}\frac{n(n-1)\cdots(n-k+1)}{n^k}\left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \lim_{n\to\infty}\frac{\lambda^k}{k!}\left(\frac{n}{n}\right)\left(\frac{n-1}{n}\right)\cdots\left(\frac{n-k+1}{n}\right)\left(1-\frac{\lambda}{n}\right)^{n-k} \qquad (\because \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x)$$

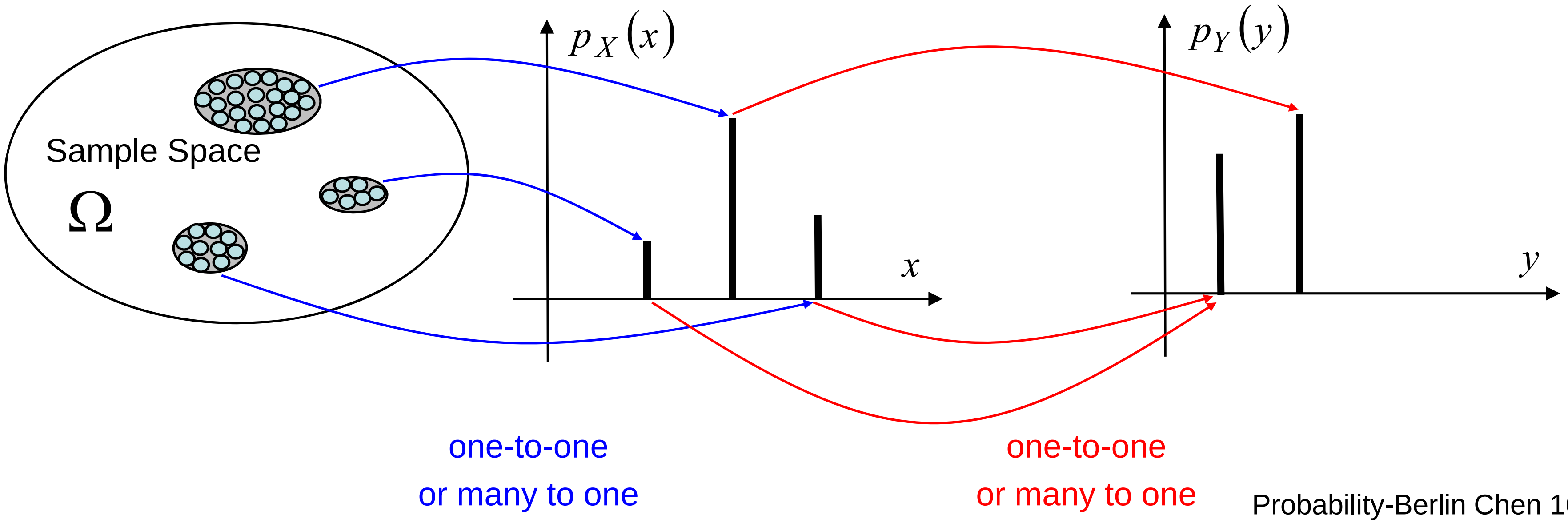
$$= \lim_{n\to\infty}\frac{\lambda^k}{k!}e^{-\lambda}$$

Functions of Random Variables (1/2)

- Given a random variable X, other random variables can be generated by applying various transformations on X
  - Linear $Y = g(X) = aX + b$

    Daily temperature in degree Fahrenheit (Y)

    Daily temperature in degree Celsius (X)
  - Nonlinear $Y = g(X) = \log X$



Functions of Random Variables (2/2)

- That is, if Y is an function of X ($Y = g(X)$), then Y is also a random variable
  - If X is discrete with PMF $p_X(x)$, then Y is also discrete and its PMF can be calculated using

$$p_Y(y) = \sum_{\{x \mid g(x) = y\}} p_X(x)$$

Functions of Random Variables: An Example

Example 2.1. Let $Y = |X|$ and let us apply the preceding formula for the PMF p_Y to the case where

$$p_X(x) = \begin{cases} 1/9 & \text{if } x \text{ is an integer in the range } [-4, 4], \\ 0 & \text{otherwise.} \end{cases}$$

The possible values of Y are $y = 0, 1, 2, 3, 4$. To compute $p_Y(y)$ for some given value y from this range, we must add $p_X(x)$ over all values x such that $|x| = y$. In particular, there is only one value of X that corresponds to $y = 0$, namely $x = 0$. Thus,

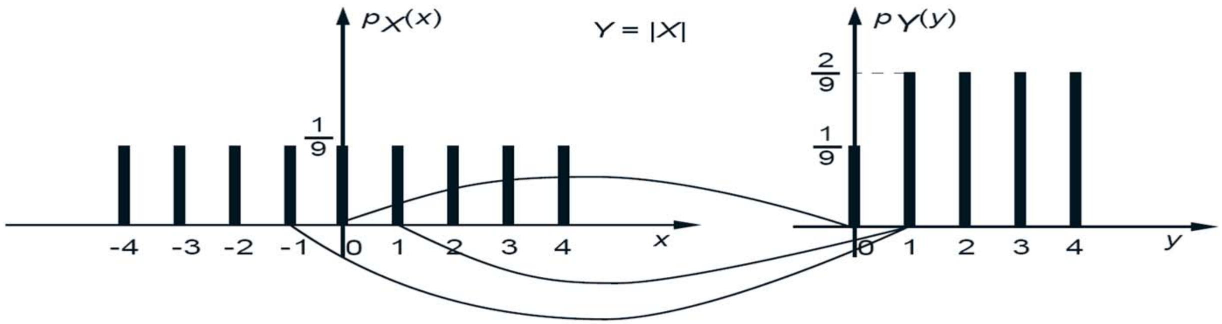
$$p_Y(0) = p_X(0) = \frac{1}{9}.$$

Also, there are two values of X that correspond to each $y = 1, 2, 3, 4$, so for example,

$$p_Y(1) = p_X(-1) + p_X(1) = \frac{2}{9}.$$

Thus, the PMF of Y is

$$p_Y(y) = \begin{cases} 2/9 & \text{if } y = 1, 2, 3, 4, \\ 1/9 & \text{if } y = 0, \\ 0 & \text{otherwise.} \end{cases}$$



Recitation

- SECTION 2.2 Probability Mass Functions
  - Problems 3, 8, 10
- SECTION 2.3 Functions of Random Variables
  - Problems 13, 14