STAT 234 Lecture 3B
Discrete Random Variables
Section 3.1-3.2 of MMSA

Yibi Huang
Department of Statistics
University of Chicago

Random Variables

Random Variables

- So far we have considered probabilities for **events** (subsets) in a space space.
- But sample spaces are often "complicated", e.g.,
  - Coin tossing: a string of outcomes such as $TTHHTTTHTHTTTTH\dots$
  - Collecting responses for a survey: a long list of the answers to all the items: (Yes;1980;3;2000$;Chicago;No;1;Maybe;N/A;7;...)
- In most cases, we are interested in some specific numerical properties computed from the "outcome" itself, e.g.,
  - # of tosses required to get the first heads
  - # of people answered `yes` to item #5 in a survey.
- Such a numerical outcome from a random phenomenon is a **random variable**.

Random Variable

Formally speaking, a **random variable** is a real-valued function on the sample space S and maps elements of S, ω, to real numbers.

$$\begin{array}{ccc} S & \xrightarrow{X} & \mathbb{R} \\ \omega & \longmapsto & x = X(\omega) \end{array}$$

Ex 1. Let X be the number of heads in 3 tosses of a coin. Sample space $S = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$. Then

$$\begin{array}{llll} X(HHH) = 3, & X(HHT) = 2, & X(HTH) = 2, & X(HTT) = 1, \\ X(THH) = 2, & X(THT) = 1, & X(TTH) = 1, & X(TTT) = 0 \end{array}$$

Ex 2. Let Y be the number of tosses required to get a head. $S = \{H, TH, TTH, TTTH, TTTTH, \ldots\}$ Then

$$Y(H) = 1,\ Y(TH) = 2,\ Y(TTH) = 3,\ Y(TTTH) = 4, \ldots$$

Discrete and Continuous Random Variable

There are two types of random variables:

- Discrete random variables can only take a finite or countable infinite number of different values
  - Example: Number of heads obtained, number of batteries replaced last year
- Continuous random variables take real (decimal) values
  - Example: lifetime of a battery, someone's blood pressure

Distribution of a Discrete Random Variable

Coin Example

Let X = number of heads in 4 tosses of a fair coin.

$$P(X = 0) = P(\{TTTT\}) = 1/16$$

Coin Example

Let X = number of heads in 4 tosses of a fair coin.

$$P(X = 0) = P(\{TTTT\}) = 1/16$$

$$P(X = 1) = P(\{HTTT, THTT, TTHT, TTTH\}) = 4/16$$

Coin Example

Let X = number of heads in 4 tosses of a fair coin.

$$P(X = 0) = P(\{TTTT\}) = 1/16$$

$$P(X = 1) = P(\{HTTT, THTT, TTHT, TTTH\}) = 4/16$$

$$P(X = 2) = P(\{HHTT, HTHT, HTTH, THHT, THTH, TTHH\}) = 6/16$$

Coin Example

Let X = number of heads in 4 tosses of a fair coin.

$$P(X = 0) = P(\{TTTT\}) = 1/16$$

$$P(X = 1) = P(\{HTTT, THTT, TTHT, TTTH\}) = 4/16$$

$$P(X = 2) = P(\{HHTT, HTHT, HTTH, THHT, THTH, TTHH\}) = 6/16$$

$$P(X = 3) = P(\{HHHT, HHTH, HTHH, THHH\}) = 4/16$$

$$P(X = 4) = P(\{HHHH\}) = 1/16$$

Coin Example

Let X = number of heads in 4 tosses of a fair coin.

$$P(X = 0) = P(\{TTTT\}) = 1/16$$

$$P(X = 1) = P(\{HTTT, THTT, TTHT, TTTH\}) = 4/16$$

$$P(X = 2) = P(\{HHTT, HTHT, HTTH, THHT, THTH, TTHH\}) = 6/16$$

$$P(X = 3) = P(\{HHHT, HHTH, HTHH, THHH\}) = 4/16$$

$$P(X = 4) = P(\{HHHH\}) = 1/16$$

The probability for each possible value of X is

Possible Value x of X	0	1	2	3	4
Probability $P(X = x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

Note: these probabilities add up to 1:

$$\frac{1}{16} + \frac{4}{16} + \frac{6}{16} + \frac{4}{16} + \frac{1}{16} = 1$$

Probability Mass Function (pmf)

Probability Mass Function (pmf)

The *probability mass function* (pmf) of a random variable X is a function $p(x)$ that maps each possible value x_i to the corresponding probability $P(X = x_i)$.

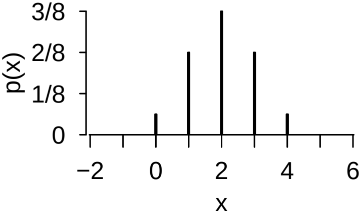
- A pmf $p(x)$ must satisfy $0 \le p(x) \le 1$ and $\sum_x p(x) = 1$.

Example (coin tossing on the previous slide)

Possible Values of X	0	1	2	3	4
Probabilities	1/16	4/16	6/16	4/16	1/16

The pmf of X is

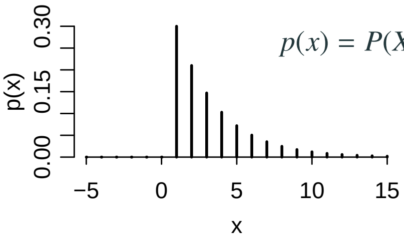
$$p(x) = \begin{cases} 1/16 & \text{if } x = 0 \text{ or } 4 \\ 4/16 & \text{if } x = 1 \text{ or } 3 \\ 6/16 & \text{if } x = 2 \\ 0 & \text{if } x \neq 0, 1, 2, 3, 4 \end{cases}$$



Example: Geometric Distribution

Let X be the number of tosses required to obtain the first heads, when tossing a coin with a probability of p to land heads.

The pmf of X is



$$
\begin{aligned}
p(x) = P(X = x) &= P(\overbrace{T \dots T}^{x-1 \text{ tails}} H) \quad \text{by indep.} \\
&= P(T)(T) \cdots P(T)P(H) \\
&= \underbrace{(1-p)(1-p)\cdots(1-p)}_{x-1 \text{ copies}}\, p \\
&= (1-p)^{x-1} p,
\end{aligned}
$$

if x is a positive integer and $p(x) = 0$ if not.

- We say X has a **geometric distribution** since the pmf is a geometric sequence
- Does $\sum_{x=1}^{\infty} p(x) = 1$?

Example: A Card Game

Consider a card game that you draw ONE card from a well-shuffled deck of cards. You win

- \$1 if you draw a heart,
- \$5 if you draw an ace (including the ace of hearts),
- \$10 if you draw the king of spades and
- \$0 for any other card you draw.

What's the pmf of your reward X?

Example: A Card Game

Consider a card game that you draw ONE card from a well-shuffled deck of cards. You win

- \$1 if you draw a heart,
- \$5 if you draw an ace (including the ace of hearts),
- \$10 if you draw the king of spades and
- \$0 for any other card you draw.

What's the pmf of your reward X?

Outcome	x	$p(x)$
Heart (not ace)	1	$12/52$
Ace	5	$4/52$
King of spades	10	$1/52$
All else	0	$35/52$

$$\Rightarrow \quad p(x) = \begin{cases} 35/52 & \text{if } x = 0 \\ 12/52 & \text{if } x = 1 \\ 4/52 & \text{if } x = 5 \\ 1/52 & \text{if } x = 10 \\ 0 & \text{for all other values of } x \end{cases}$$