# Discrete Mathematics and Its Applications 2 (CS147)

***Lecture 10: Random variable, coupon collector's problem***

**Fanghui Liu**

Department of Computer Science, University of Warwick, UK

# Target of discrete probability in this module...

**Problem (Coupon collector's problem)**

*We repeatedly sample from a set of $n$ distinct coupons until at least one copy of each distinct coupon is obtained. What is the* **expected** *times do we need?*

# Target of discrete probability in this module...

**Problem (Coupon collector's problem)**

*We repeatedly sample from a set of $n$ distinct coupons until at least one copy of each distinct coupon is obtained. What is the* **expected** *times do we need?*

- randomness, probability space
- indicator random variable
- expectation
- tail

# Random variable

A random variable (r.v.) is any rule (i.e., function) that associates a number with each outcome in the sample space.

**Definition (Random variable is a function!)**

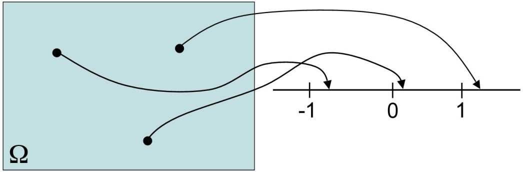
Given a probability space $(\Omega, \mathcal{F}, \Pr)$ and a function $X : \Omega \to \mathbb{R}$, if for any $a \in \mathbb{R}$, we have $\{\omega : X(\omega) \leq a\} \in \mathcal{F}$, then $X$ is a random variable.

# Random variable

A random variable (r.v.) is any rule (i.e., function) that associates a number with each outcome in the sample space.

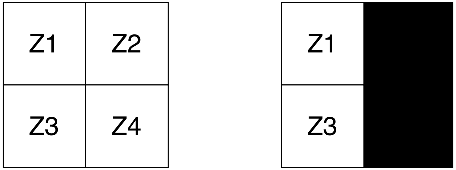
## Definition (Random variable is a function!)

Given a probability space $(\Omega, \mathcal{F}, \Pr)$ and a function $X : \Omega \to \mathbb{R}$, if for any $a \in \mathbb{R}$, we have $\{\omega : X(\omega) \leq a\} \in \mathcal{F}$, then $X$ is a random variable.



# *Recall our example in Lecture 8...

- case 1: a transparent box (left)
- case 2: half covered by opaque cloth



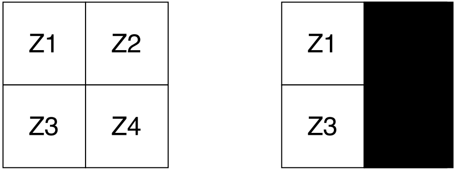

sample space $\Omega = \{Z1, Z2, Z3, Z4\}$

- case 1: $\mathcal{F}_1$ : a collection of all subsets of $\Omega$
- case 2: $\mathcal{F}_2$ is

$$\mathcal{F}_2 = \{\Omega, \emptyset, \{Z1\}, \{Z2, Z3, Z4\}, \\ \{Z3\}, \{Z1, Z2, Z4\}, \{Z1, Z3\}, \{Z2, Z4\}\}$$

# *Recall our example in Lecture 8...

- case 1: a transparent box (left)
- case 2: half covered by opaque cloth




Define a function $X : \Omega \to \mathbb{R}$ as

$$X(\omega) = \begin{cases} 1, \text{ if } \omega = Z1 \\ 1.6, \text{ if } \omega = Z2 \\ 4.3, \text{ if } \omega = Z3 \\ 5, \text{ if } \omega = Z4 \end{cases}$$

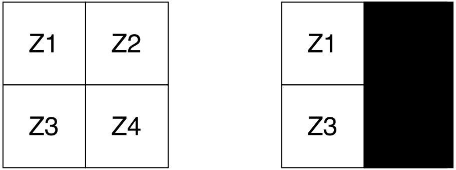
sample space $\Omega = \{Z1, Z2, Z3, Z4\}$

- case 1: $\mathcal{F}_1$ : a collection of all subsets of $\Omega$
- case 2: $\mathcal{F}_2$ is

$$\mathcal{F}_2 = \{\Omega, \emptyset, \{Z1\}, \{Z2, Z3, Z4\}, \{Z3\}, \{Z1, Z2, Z4\}, \{Z1, Z3\}, \{Z2, Z4\}\}$$

# *Recall our example in Lecture 8...

- case 1: a transparent box (left)
- case 2: half covered by opaque cloth




Define a function $X : \Omega \to \mathbb{R}$ as

$$X(\omega) = \begin{cases} 1, \text{ if } \omega = Z1 \\ 1.6, \text{ if } \omega = Z2 \\ 4.3, \text{ if } \omega = Z3 \\ 5, \text{ if } \omega = Z4 \end{cases}$$

$X$ is a random variable w.r.t $\mathcal{F}_1$ but not a random variable w.r.t $\mathcal{F}_2$ because

$$\{\omega : X(\omega) \leq 2\} = \{Z1, Z2\} \notin \mathcal{F}_2 \,.$$

sample space $\Omega = \{Z1, Z2, Z3, Z4\}$

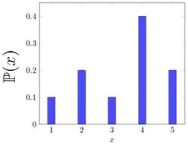
- case 1: $\mathcal{F}_1$ : a collection of all subsets of $\Omega$
- case 2: $\mathcal{F}_2$ is

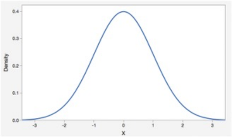
$$\mathcal{F}_2 = \{\Omega, \emptyset, \{Z1\}, \{Z2, Z3, Z4\}, \\ \{Z3\}, \{Z1, Z2, Z4\}, \{Z1, Z3\}, \{Z2, Z4\}\}$$

# Types of random variables

◦ A random variable (r.v.) can be either discrete or continuous

- discrete r.v.: has a countable number of possible values
- continuous r.v.: takes all values in an interval of numbers



In this module, we mainly consider discrete random variables

# Indicator function

## Definition

Let $A \subseteq \Omega$, define

$$1_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{otherwise} \end{cases}$$

transform operations of sets into algebra operations!

## Statement

$$A = B \Leftrightarrow 1_A = 1_B$$

$$A \subseteq B \Leftrightarrow 1_A \leq 1_B$$

$$1_{A \cap B} = \min\{1_A, 1_B\} = 1_A 1_B$$

$$1_{A \cup B} = \max\{1_A, 1_B\}$$

# Indicator function

## Definition

Let $A \subseteq \Omega$, define

$$1_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{otherwise} \end{cases}$$

transform operations of sets into algebra operations!

## Statement

$$A = B \Leftrightarrow 1_A = 1_B$$
$$A \subseteq B \Leftrightarrow 1_A \leq 1_B$$
$$1_{A \cap B} = \min\{1_A, 1_B\} = 1_A 1_B$$
$$1_{A \cup B} = \max\{1_A, 1_B\}$$

## Example (in sorting algorithms)

For an array with size $n$, denote a random variable $Y_{ij}$ with $i, j \in [n]$ as

$$Y_{ij} = \begin{cases} 1 & \text{if } a_i, a_j \text{ are compared} \\ 0 & \text{otherwise.} \end{cases}$$

# Probability Mass Function

## Definition

The probability mass function (PMF) of a **discrete** random variable is defined as

$$f_X(a) = \Pr(X = a) = \Pr(\{\omega \in \Omega : X(\omega) = a\}).$$

**Remark:** $\sum_{a \in X(\omega)} f_X(a) = 1$.

# Probability Mass Function

## Definition

The probability mass function (PMF) of a **discrete** random variable is defined as

$$f_X(a) = \Pr(X = a) = \Pr(\{\omega \in \Omega : X(\omega) = a\}) .$$

**Remark:** $\sum_{a \in X(\omega)} f_X(a) = 1$.

## Example

Consider a biased coin flipped with $p$ for head, $1 - p$ for the tail, we denote

| $a$ | $1$ | $0$ |
|---|---|---|
| $\Pr[X = a]$ | $p$ | $1 - p$ |

# Distribution of a random variable

The distribution of a random variable describes the probability that it takes on various values.

## Definition (Cumulative distribution function)

For real-valued random variables, the distribution function or cumulative distribution function is a function $F_X(a) = \Pr(X \leq a)$

## Distribution of a random variable

The distribution of a random variable describes the probability that it takes on various values.

### Definition (Cumulative distribution function)

For real-valued random variables, the distribution function or cumulative distribution function is a function $F_X(a) = \Pr(X \leq a)$

**Remark:** for discrete random variables: take on only countably many possible values $a_1, a_2, \ldots, a_n, \ldots$.

| $X(\omega)$ | $a_1$ | $a_2$ | $\ldots$ | $a_i$ | $\ldots$ |
|---|---|---|---|---|---|
| probability | $\Pr(X = a_1)$ | $\Pr(X = a_2)$ | $\ldots$ | $\Pr(X = a_i)$ | $\ldots$ |

## Distribution of a random variable

The distribution of a random variable describes the probability that it takes on various values.

### Definition (Cumulative distribution function)

For real-valued random variables, the distribution function or cumulative distribution function is a function $F_X(a) = \Pr(X \leq a)$

**Remark:** for discrete random variables: take on only countably many possible values $a_1, a_2, \ldots, a_n, \ldots$.

| $X(\omega)$ | $a_1$ | $a_2$ | $\ldots$ | $a_i$ | $\ldots$ |
|---|---|---|---|---|---|
| probability | $\Pr(X = a_1)$ | $\Pr(X = a_2)$ | $\ldots$ | $\Pr(X = a_i)$ | $\ldots$ |

### Property

- $F(-\infty) = 0, F(\infty) = 1$ *and* $F$ *is non-decreasing*
- $\Pr(a < X \leq b) = F_X(b) - F_X(a)$
- *right-continuous* $F_X(a^+) = F_X(a)$

# Example: typical discrete random variables

- Bernoulli distribution: $X \sim \text{Bernoulli}(p)$
  $\Pr(X = 1) = p$ and $\Pr(X = 0) = 1 - p$.

# Example: typical discrete random variables

- Bernoulli distribution: $X \sim$ Bernoulli$(p)$
  $\Pr(X = 1) = p$ and $\Pr(X = 0) = 1 - p$.
- Binomial distribution: $X \sim$ Binomial$(n, p)$
  $\Pr(X = k) = \binom{n}{k} p^k q^{(n-k)}$, where $n$ and $p$ are parameters of the distribution and $q = 1 - p$
  - Experiment consists of $n$ trials
  - Trials are identical and independent
  - Constant probability for each observation

## Example: typical discrete random variables

- Bernoulli distribution: $X \sim \text{Bernoulli}(p)$
  $\Pr(X = 1) = p$ and $\Pr(X = 0) = 1 - p$.
- Binomial distribution: $X \sim \text{Binomial}(n, p)$
  $\Pr(X = k) = \binom{n}{k} p^k q^{(n-k)}$, where $n$ and $p$ are parameters of the distribution and $q = 1 - p$
  - Experiment consists of $n$ trials
  - Trials are identical and independent
  - Constant probability for each observation

### Geometric distribution

$X \sim \text{Geo}(p) : \Pr(X = k) = (1 - p)^{k-1} p, \forall k \geq 1.$

◦ number of tails we flip before we get the first head in a sequence of biased coin-flips.

- Fails in the first $k - 1$ times

$$\Pr(X = k) = (1 - p)^{k-1} \; p$$

- Success at the $k$-th time

# Geometric distribution in coupon collector's problem

## Problem

*We repeatedly sample from a set of $n$ distinct coupons until at least one copy of each distinct coupon is obtained. What is the* **expected** *times do we need?*

# Geometric distribution in coupon collector's problem

## Problem

*We repeatedly sample from a set of $n$ distinct coupons until at least one copy of each distinct coupon is obtained. What is the* **expected** *times do we need?*

- Sample a new object: ✓
- Sample a repeated object: ✗

$\Rightarrow$ success to sample a new object before previous (repeated) objects

# Geometric distribution in coupon collector's problem

## Problem

*We repeatedly sample from a set of $n$ distinct coupons until at least one copy of each distinct coupon is obtained. What is the* **expected** *times do we need?*

- Sample a new object: ✓
- Sample a repeated object: ✗

$\Rightarrow$ success to sample a new object before previous (repeated) objects

- $\Pr(\text{find the first unique coupon}) = \frac{n}{n} = 1$
- $\Pr(\text{find the second unique coupon}) = \frac{n-1}{n}$
- $\Pr(\text{find the } i\text{-th unique coupon}) = \frac{n-(i-1)}{n}$

# Geometric distribution in coupon collector's problem

## Problem

*We repeatedly sample from a set of $n$ distinct coupons until at least one copy of each distinct coupon is obtained. What is the* **expected** *times do we need?*

- Sample a new object: ✓
- Sample a repeated object: ✗

$\Rightarrow$ success to sample a new object before previous (repeated) objects

- $\Pr(\text{find the first unique coupon}) = \frac{n}{n} = 1$
- $\Pr(\text{find the second unique coupon}) = \frac{n-1}{n}$
- $\Pr(\text{find the } i\text{-th unique coupon}) = \frac{n-(i-1)}{n}$

Let $t_i$ = times to collect the $i$-th unique coupon after collecting $(i-1)$-th unique coupons.

$\Rightarrow t_i \sim \text{Geo}(\frac{n-(i-1)}{n})$.

# Continuous random variable

## Definition

A random variable is continuous if $\Pr(X = a) = 0$ for any $a \in \mathbb{R}$.

**Remark:** not assign probability to points but to **intervals**

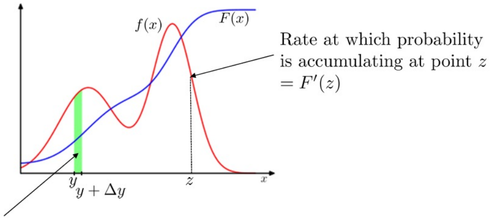
# Continuous random variable

## Definition

A random variable is continuous if $\Pr(X = a) = 0$ for any $a \in \mathbb{R}$.

**Remark:** not assign probability to points but to **intervals**

- A probability density function (PDF) $f_X$: $\Pr(a \leq X \leq b) = \int_a^b f_X(x)\mathrm{d}x$.
- A cumulative distribution function (CDF) is $F_X(a) = \int_{-\infty}^a f_X(t)\mathrm{d}t$.




$$P(y < X \leq y + \Delta y) = F(y + \Delta y) - F(y)$$
$$\approx f(y)\Delta y \text{ as } \Delta y \to 0$$

(also called AREA under the pdf)

# Example: typical continuous random variables

- Uniform distribution over $[a, b]$:

$$F_X(x) = \begin{cases} 0, \text{ if } x \leqslant a \\ (x-a)/(b-a), \text{ if } a < x < b \\ 1, \text{ if } x > b \end{cases}$$

- Gaussian distribution ("most important" distribution in probability theory):

$$F_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \mathrm{d}x \,.$$

# Joint distribution

Two or more (discrete) random variables can be described using a joint distribution, which can be represented as $\Pr[X = x, Y = y]$ for two random variable $X$ and $Y$
Marginal distribution:

$$\Pr(X = x) = \sum_{y} \Pr[X = x, Y = y]$$

**Example**

Let $X$ and $Y$ be six-sided dices, then $\Pr[X = x, Y = y] = 1/36$ for all values $x$ and $y$ in $\{1,2,3,4,5,6\}$

# Independent Random Variables

## Definition

Two discrete random variables $X$ and $Y$ over $(\Omega, \mathcal{F}, \Pr)$ are said to be independent if and only if for every $x$ in the range of $X$ and $y$ in the range of $Y$

$$\Pr[(X = x) \cap (Y = y)] = \Pr(X = x) \cdot \Pr(Y = y)$$