Key concepts for COMP1215 and reading list

Andersen Ang @ ECS, University of Southampton andersen.ang@soton.ac.uk version: December 9, 2023

(For the parts on combinatorics, graph, probability and statistics of COMP1215, October semester of the year 2023)

1 Key concepts in combinatorics

1.1 Basic counting techniques from set theory

1. Given a set A, the number of elements of A is denoted as $|A|$
2. Inclusion-exclusion principle $|A \cup B| = |A| + |B| - |A \cap B|$
   Its generalisation to 3 sets.
3. Counting by enumeration: list all possible outcome.
   Draw tree to visualize.
4. Sum rule as a special case of inclusion-exclusion principle: $|A \cup B| = |A| + |B|$ if A, B disjoint
5. Product rule as application of Cartesian product $|A \times B| = |A| \cdot |B|$
6. Subtraction rule as application of complement and disjoint set: If $A \subset S$ then for $A^c = S \setminus A$, we have $|A^c| = |S| - |A|$
7. Floor function $\lfloor x \rfloor \coloneqq \max\{m \in \mathbb{Z} \mid m \leq x\}$, and counting number of divisible integers

1.2 Basic counting techniques from binomial coefficient

1. Factorial $k! = \begin{cases} 1 & k = 0 \\ k(k-1)! & k > 0 \end{cases}$
2. binomial coefficient $\binom{n}{k} = \dfrac{n!}{(n-k)!k!}$
3. trinomial coefficient $\binom{n}{i,j,k} = \dfrac{n!}{i!j!k!}$ and multinomial coefficient $\binom{n}{k_1, \cdots, k_r} = \dfrac{n!}{k_1! \cdots k_r!}$
4. Definition of permutation: an ordered arrangement of a set of distinct objects
5. Definition of combination: an unordered arrangement of a set of distinct objects
6. $n!$ = number of permutation of n-set of distinct object
7. $\binom{n}{k}$ = number of n-choose-k combination of a n-set of distinct objects
8. $\binom{n}{k_1, k_2, ..., k_r}$ = number of permutation of a n-set with $(k_1, k_2, ..., k_r)$ repeated object
9. $\binom{n+k-1}{k}$ = number of n-choose-k combination of a n-set with k repeated objects

	order important	order not important
no repetition	permutation $n!$	combination $\binom{n}{k}$
repetition	generalized permutation $\binom{n}{n_1, n_2 \ldots, n_r}$	generalized combination $\binom{n+k-1}{k}$

10. Number of shortest lattice paths from $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ to $\begin{bmatrix} m \\ n \end{bmatrix}$ is $\binom{m+n}{m}$
11. Binomial expansion $(x+y)^n$
12. Solving trinomial expansion using binomial expansion
13. Properties of binomial coefficient, Pascal's triangle and Fibonacci sequence

1.3 Advanced counting techniques

- Counting by bijection: Iif counting directly on X is difficult
  - construct a function $x \mapsto f(x)$ that $f : X \to Y$ is bijective
  - count Y instead
- Generating function technique $G(x)$ in counting and recursion
  - Counting by polynomial coefficient $[x^n]$ on the outcome of tossing dice
  - Geometric series: $1 + x + x^2 + ... + x^n = \dfrac{1 - x^{n+1}}{1 - x}$ for $x \neq 1$ and $1 + x + x^2 + \cdots = \dfrac{1}{1 - x}$ for $|x| < 1$
  - $\Big(1 + x + x^2 + ...\Big)^n = \sum_{r=0}^{\infty} \binom{n + r - 1}{r} x^r$ (this formula will be given in the question if needed)
  - Partial fraction decomposition of $\dfrac{f(x)}{g(x)}$
- Pigeonhole principle

2 Key concepts in basic graph theory

1. Representation of a graph: VVEEMAD
   - V is the set of nodes
   - E is the set of edge (i, j)
   - Incidence matrix $M_{ij} = 1$ if $(i, j) \in E$
   - Adjacency matrix $A_{ij} = 1$ if $(i, j) \in E$ or $(j, i) \in E$
   - D_{ii} is how many edge touching node i
2. Not in exam
   - Directed vs undirected, weighted, simple vs multi-edge
   - Handshaking lemma, all graph has even number of odd vertices
   - Null graph, complete graph
   - Walk, trail, path, cycle, tree, forest, bipartite
   - Number of k-walks in a graph
   - Graph coloring and chromatic polynomial

3 Key concepts in probability

3.1 Classical probability from set theory

1. Sample space Ω, event E and probability $\mathbb{P}(E) := \dfrac{|E|}{|\Omega|}$
2. The probability axioms
   - Axiom-0 $\Omega \neq \varnothing$
   - Axiom-1 $\mathbb{P}(E) \geq 0$ for any event $E \subset \Omega$
   - Axiom-2 $\mathbb{P}(\Omega) = 1$ for any Ω
   - Axiom-3 Sum rule for disjoint event if $E \subset \Omega$, $F \subset \Omega$ and $E \cap F = \varnothing$, then $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F)$
3. Complementary event, mutually exclusive events
4. Using combinatorics in the classical probability of tossing multiple dices

3.2 Random variable

1. Joint probability $\mathbb{P}(X = x, Y = y)$ means the probability of the event $\{X = x \text{ AND } Y = y\}$
2. Marginal probability $\mathbb{P}(X = x) = \sum_{y \in \mathcal{Y}} \mathbb{P}(X = x, Y = y)$
3. Conditional probability $\mathbb{P}(X = x|Y = y) := \dfrac{\mathbb{P}(X = x, Y = y)}{\mathbb{P}(Y = y)}, \ \mathbb{P}(Y = y) \neq 0$
4. Independent random variables $\mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$
5. Expected value $\mathbb{E}[X] := \sum_{x \in \mathcal{X}} xp(x)$ and $\mathbb{E}[f(X)] := \sum_{x \in \mathcal{X}} f(x)p(x)$, where $p(x) := \mathbb{P}(X = x)$
6. Expectation is linear $\mathbb{E}\Big[aX + bY + c\Big] = a\mathbb{E}[X] + b\mathbb{E}[Y] + c$
7. Variance $\mathbb{V}[X] := \mathbb{E}\Big[\big(X - \mathbb{E}[X]\big)^2\Big] = \mathbb{E}[X^2] - \Big(\mathbb{E}[X]\Big)^2$
8. Variance quadratic formula $\mathbb{V}\Big[aX \pm bY + c\Big] = a^2\mathbb{V}[X] \pm 2ab\,\mathrm{cov}(X, Y) + b^2\mathbb{V}[Y]$
9. $\mathrm{cov}(X, Y) := \mathbb{E}\Big[\big(X - \mathbb{E}[X]\big)\big(Y - \mathbb{E}[Y]\big)\Big]$
   how to remember: replace one X in $\mathbb{V}[X]$ by Y
10. $\mathrm{corr}(X, Y) := \dfrac{\mathrm{cov}(X, Y)}{\sqrt{\mathbb{V}[X]\mathbb{V}[Y]}}$ as normalized covariance
11. Distinguish between $\mathbb{E}[X + Y], \mathbb{E}[XY]$ and $\mathbb{E}[(X, Y)]$

4 Key concepts in statistics

4.1 Probability distribution function (Not in exam 2023)

1. Probability = area under the curve of PDF
2. Crazy things about probability for continuous random variable
   (a) $\mathbb{P}(X \leq a) = \mathbb{P}(X < a)$ because $\mathbb{P}(X = a) = 0$ this is not true for discrete random variable
   (b) Zero probability $\neq$ impossible to occur
3. Bernoulli distribution
   - $\mathcal{X} = \{0, 1\}$
   - $X \sim \mathrm{Ber}(\theta)$ then $\mathbb{P}(X = x|\theta) = p(x|\theta) = \theta^x(1 - \theta)^{1-x}$
   - It is used for modelling binary event
4. Binomial distribution
   - $\mathcal{X} = \{0, 1, ..., n\}$
   - $X \sim \mathrm{Bin}(n, \theta)$ then $\mathbb{P}(X = x|n, \theta) = p(x|n, \theta) = \binom{n}{x}\theta^x(1 - \theta)^{n-x}$
   - It is used for modelling number of success in n binary event
5. Poisson distribution
   - $\mathcal{X} = \{0, 1, 2, ...\}$
   - $X \sim \mathrm{Poi}(\lambda)$ then $\mathbb{P}(X = x|\lambda) = p(x|\lambda) = \dfrac{\lambda^x e^-\lambda}{x!}$
   - It is used for modelling time interval event

4.2 Normal random variable (important)

1. Normal random variable
   - $\mathcal{X} = \mathbb{R}$
   - $X \sim \mathcal{N}(\mu, \sigma^2)$ then $\mathbb{P}(X = x|\mu, \sigma^2) = p(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, it means "the probability of X takes the value x, given that X is a normal random variable with parameter μ, σ^2, has the expression $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$"
   - $\mathbb{E}[X] = \mu$, the expected value of X is μ
   - $\mathbb{V}[X] = \sigma^2$, the variance of X is σ^2
   - The curve $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ is symmetric (even function)
2. Standard Normal random variable $Z \sim \mathcal{N}(0, 1)$ and $\mathbb{P}(Z = z|0, 1) = p(z|0, 1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$
3. If $Z \sim \mathcal{N}(0, 1)$ then $\mathbb{P}(a \le Z \le b) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = \frac{\text{erf}(b) - \text{erf}(a)}{2}$
4. The error function $\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$ can only be calculated by computer
5. Standardization: if $X \sim \mathcal{N}(\mu, \sigma^2)$ then $Z = \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$

4.3 Point estimation

- Sample mean / average $\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$, the symbol $\overline{x}$ is pronounced as "x-bar"
- **Theorem** If $X_1, ..., X_n \sim \mathcal{N}(\mu, \sigma^2)$ then sample mean $\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$
- Unbiased estimator of variance $s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$
  - Sample variance $\frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2$ is a biased estimator

4.4 Confidence interval

1. The $100(1 - \alpha)\%$ confidence interval of population mean, with known variance is

$$T_\alpha = \left[\overline{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}},\ \overline{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}\right]$$

   where $Z \sim \mathcal{N}(0, 1)$. For 95% confidence interval, $\alpha = 0.05$ and $z_{\alpha/2} = 1.96$ (computed by solving error function).

2. (Not in exam) The $100(1 - \alpha)\%$ confidence interval of population mean, with unknown variance is

$$T_\alpha = \left[\overline{x} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}},\ \overline{x} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}\right]$$

   where $T \sim T(n - 1)$ is the T distribution with $n - 1$ degree of freedom. The value $t_{\alpha/2, n-1}$ is computed by computer.

3. (Not in exam) The $100(1 - \alpha)\%$ confidence interval of difference of two population mean, with known variance is

$$T_\alpha = \left[\overline{x}_A - \overline{x}_B - z_{\alpha/2} \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}},\ \overline{x}_A - \overline{x}_B + z_{\alpha/2} \sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}\right]$$

   we use the same formula for the case of unknown variance, by replacing σ^2 with s^2

4.5 Hypothesis testing

1. We take the null hypothesis (innocent) as the default position, and we use data to get evidence against the null

2. p-value is the probability $\mathbb{P}(\text{data}|H_0)$

   - $\mathbb{P}(\text{data}|H_0)$ means "the conditional probability of seeing the data given that the null hypothesis H_0 is true"
   - small p-value means we have lots of evidence against the null, so null is probably false
   - large p-value means we do not have enough evidence against the null, no conclusion
     Note that a large p-value does not mean H_0 is true

3. For $\begin{array}{ccc} H_0 & : & \mu = \mu_0 \\ & \text{vs} & \\ H_A & : & \mu \neq \mu_0 \end{array}$, the p-value is

$$p = 2\mathbb{P}\Big(Z < -|z^*|\Big) = 2\int_{-\infty}^{z^*} p(z)dz \quad \text{where} \quad z^* = \frac{\overline{x} - \mu_0}{\sigma/\sqrt{n}}$$

   $\overline{x}$ is the sample mean / data average, μ_0 is the guess value in the null hypothesis (you make the hypothesis that the unknown μ is μ_0), σ is the standard deviation and n is the number of data point

   If σ is unknown we estimate it using the unbiased estimator of variance $s = \sqrt{\dfrac{1}{n-1}\displaystyle\sum_{i=1}^{n}(x_i - \overline{x})^2}$

5 So, you want a reading list?

- **Discrete Mathematics and Its Applications, 7th Edition** Author: Kenneth Rosen
  QA39 ROS in Hartley Library
  - Chapter 6 Counting page 385-444
  - Chapter 7 Discrete Probability page 445-494
    Skip Section 7.3
  - Chapter 8 Advanced Counting Techniques
    Section 8.4 Generating Functions page 537-552
    Section 8.5 Inclusion-Exclusion page 552-558
  - Chapter 10 Graphs page 641-678

  Older or newer version are also ok

- **Introductory Statistics, 3rd edition** Author: Sheldon Ross
  The library has the ebook (click "view ebook" on the right)
  Basically just read the first 500 pages
  - Chapter 1-3 are "high school level" you can skip or read
  - Chapter 4 Probability you can skip or read
  - Chapter 5 Discrete Random Variables page 209-page 260
    Focus on 5.1-5.5
  - Chapter 7 Distributions of Sampling Statistics you can skip or read
    You can read the whole chapter, skip the whole chapter, or just focus on 7.3 and 7.4
  - Chapter 6 Normal Random Variables page 261-296
  - Chapter 8.5-8.7 page 347-386
  - Chapter 9 Testing Statistical Hypotheses page 387-442
    Focus on 9.1-9.4
  - Chapter 10 Hypothesis Tests Concerning Two Populations page 443-502
    Focus on 10.1-10.3

- **Foundations of Computer Science, 2nd Revised Edition** Compilded by Powel Sobocinski
  QA39 SOB in Hartley library, 36 copies
  - Chapter 6.1-6.3 page 505-539
  - Chapter 6.5-6.8 page 545-569
  - Chapter 8.1-8.2 page 617-640

- Chapter 8.5-8.6 page 653-665
- Chapter 9.1-9.2 page 681-694

- **Schaum's outlines Discrete Mathematics, 3rd edition** Author: Seymour Lipschutz
  QA43 LIP in Hartley library, have ebook
  - Chapter 5 to Chapter 9 page 88-204

6 Advanced / challenging books

- **Concrete Mathematics A foundation for Computer Science, 2nd edition** author: Graham, Knuth, Patashnik
  QA39 GRA in Hartley library
  - Chapter 2 Sums page 21-25, page 30-41
  - Chapter 3.1 Floors and Ceilings page 67-78
  - Chapter 5 Binomial Coefficients
    * Ch5.1 page 153-172
    * Ch5.4 page 196-204
  - Chapter 7 Generating Functions page 320-380
  - Chapter 8 Discrete Probability page 381-38
- **Counting : the art of enumerative combinatorics** author: George E Martin
  QA164.8 MAR

7 Very challenging books

- **102 Combinatorial Problems: From the Training of the USA IMO Team**
- **Enumerative combinatorics. Volume 1** Author: Richard P Stanley
  ebook avaliable in the library

END OF PDF