REPORT

February 2, 2025

1 Information Theory CheatSheet

1.1 Quick Review

1.1.1 Information Theory CheatSheet

(Based on Elements of Information Theory, 2nd Edition by Thomas M. Cover, Joy A. Thomas)

---

1. Capacity Regions

- **Multiple Access Channel (MAC):**
  Capacity region:
  $R_1 \leq I(X_1; Y|X_2), \quad R_2 \leq I(X_2; Y|X_1), \quad R_1 + R_2 \leq I(X_1, X_2; Y)$
- **Broadcast Channel:**
  - No general formula for all cases.
  - For **degraded channels**, optimal rates achieved using **superposition coding**:
    $R_1 \leq I(X; Y_1), \quad R_2 \leq I(X; Y_2|Y_1)$

---

2. Markov Chains

- **Definition:**
  A stochastic process where future states depend only on the current state:
  $P(X_{n+1}|X_n, X_{n-1}, \dots) = P(X_{n+1}|X_n)$
- **Entropy Rate:**
  $H(X) = \lim_{n\to\infty} \frac{H(X_1, X_2, \dots, X_n)}{n}$
- **Stationary Distribution:**
  For transition matrix P, solve $\pi P = \pi$.

---

3. Maximization of Entropy

- **Discrete case:**
  Entropy is maximized when all outcomes are equally likely:
  $H(X) \leq \log_2 |\mathcal{X}|$

- **Continuous case:**
  Differential entropy is maximized by a Gaussian distribution:
  $h(X) \leq \frac{1}{2}\log_2(2\pi e\sigma^2)$

---

4. Capacities of Different Channels

1. **Binary Symmetric Channel (BSC):**
   $C = 1 - H(p), \quad H(p) = -p\log_2 p - (1-p)\log_2(1-p)$
2. **Binary Erasure Channel (BEC):**
   $C = 1 - p$
3. **AWGN Channel:**
   $C = \frac{1}{2}\log_2\left(1 + \frac{P}{N_0 B}\right)$

---

5. Calculate Entropy of Channels

- **Mutual Information:**
  $I(X;Y) = H(Y) - H(Y|X)$
- **Entropy of a channel with output Y:**
  $H(Y) = -\sum_{y \in \mathcal{Y}} P(y)\log_2 P(y)$

---

6. Index Coding

- **Definition:**
  Reduce the number of transmissions by using side information at clients.
- **Example:**
  For messages W_1, W_2, W_3 and side information:
  - Client 1 knows W_2
  - Client 2 knows W_3
  - Client 3 knows W_1
    Optimal coded transmissions: $W_1 \oplus W_2, W_2 \oplus W_3, W_3 \oplus W_1$.

---

7. Network Coding

- **Definition:**
  Intermediate nodes perform operations (e.g., XOR) on data streams to increase throughput.
- **Example:**
  In a butterfly network, transmit $X = A \oplus B$. Both sinks decode:
  $A = X \oplus B, \quad B = X \oplus A$

---

8. Coded Caching

- **Basic Idea:**
  Pre-store coded data at users to reduce peak-time traffic.
- **Formula:**
  $L = \frac{N(1-M/N)}{1+KM/N}$ where N is the number of files, M is the cache size per user, and K is the number of users.

---

9. Gambling (after 10 goals)

- **Kelly Criterion:**
  Maximizes logarithmic utility by choosing the optimal bet fraction:
  $f^* = \frac{bp-q}{b}, \quad q = 1-p$
- **Example:**
  If $p = 0.6$ and odds $b = 2$, the optimal bet is:
  $f^* = \frac{2\cdot0.6-0.4}{2} = 0.4$

---

10. MAC or Broadcast Channel (Optimal Schemes)

- **MAC:**
  Achieve optimal rates using **successive interference cancellation**:
  $R_1 \leq I(X_1; Y|X_2), \quad R_2 \leq I(X_2; Y|X_1)$
- **Broadcast:**
  Achieve capacity using **superposition coding**:
  $X = \alpha X_1 + (1-\alpha)X_2$

---

11. Asymptotic Equipartition Property (AEP)

- **Definition:**
  For a sequence of i.i.d. random variables, the probability of typical sequences converges to:
  $P(x^n) \approx 2^{-nH(X)}$
- **Implications:**
  - Most sequences are typical as $n \to \infty$.
  - Supports **data compression** and **channel coding** by focusing on typical sequences.

---

12. Coded MapReduce

- **Definition:**
  Encode intermediate data to reduce communication during the shuffle phase.
- **Example:**
  If there are 4 mappers and 3 reducers, coded transmissions allow each reducer to decode its required data from fewer transmissions.
- **Communication Reduction:**
  $R = \frac{1}{r}$ where r is the number of reducers.

This cheat sheet covers essential formulas, examples, and definitions for each topic, providing a quick reference for **Information Theory** concepts.

1.2 Exercices 1

1.2.1 Information Theory Q&A with Mathematical Problems

(With focus on AEP and related concepts)

---

1. Capacity Regions

1. **Q:**
   For a two-user Gaussian multiple access channel (MAC) with $P_1 = 4$, $P_2 = 6$, and noise $N = 2$, find the sum-rate constraint.
   A:
   The sum-rate constraint is:
   $R_1 + R_2 \le \frac{1}{2}\log_2\left(1 + \frac{P_1+P_2}{N}\right) = \frac{1}{2}\log_2\left(1 + \frac{4+6}{2}\right) = 1.8\,\text{bits}$
2. **Q:**
   Explain how the capacity region changes when time-sharing is used in a broadcast channel.
   A:
   Time-sharing allows convex combinations of achievable rate points, expanding the capacity region by alternating between different transmission schemes.

---

2. Markov Chains

1. **Q:**
   For a Markov chain with transition matrix
   $P = \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix}$, find the stationary distribution.
   A:
   Solve $\pi P = \pi$ with $\pi_1 + \pi_2 = 1$:
   $\pi_1 = 0.6\pi_1 + 0.3\pi_2, \quad \pi_2 = 0.4\pi_1 + 0.7\pi_2$ Solution: $\pi = (0.43, 0.57)$.
2. **Q:**
   Calculate the entropy rate of this Markov chain.

A:
$H(X) = \sum_{i,j} \pi(i) P_{ij} \log_2 \frac{1}{P_{ij}}$ Substituting values:
$H(X) = 0.43 \cdot (0.6 \log_2 \frac{1}{0.6} + 0.4 \log_2 \frac{1}{0.4}) + 0.57 \cdot (0.3 \log_2 \frac{1}{0.3} + 0.7 \log_2 \frac{1}{0.7})$

---

3. Maximization of Entropy

1. **Q:**
   Prove that entropy is maximized for a discrete variable when all outcomes are equally likely.
   A:
   If $p(x) = \frac{1}{|\mathcal{X}|}$, then:
   $H(X) = -\sum_{x \in \mathcal{X}} \frac{1}{|\mathcal{X}|} \log_2 \frac{1}{|\mathcal{X}|} = \log_2 |\mathcal{X}|$

2. **Q:**
   Calculate the differential entropy of a Gaussian variable with variance $\sigma^2 = 3$.
   A:
   $h(X) = \frac{1}{2} \log_2(2\pi e \sigma^2) = \frac{1}{2} \log_2(2\pi e \cdot 3) \approx 2.77\,\text{bits}$

---

4. Capacities of Different Channels

1. **Q:**
   For a binary symmetric channel (BSC) with $p = 0.2$, calculate the channel capacity.
   A:
   $C = 1 - H(p), \quad H(p) = -p \log_2 p - (1-p) \log_2(1-p)$ Substitution gives $H(p) \approx 0.72$, so $C \approx 0.28\,\text{bits}$.

2. **Q:**
   Find the capacity of an AWGN channel with power $P = 10$, noise spectral density $N_0 = 1$, and bandwidth $B = 1$.
   A:
   $C = \frac{1}{2} \log_2 \left(1 + \frac{P}{N_0 B}\right) = \frac{1}{2} \log_2(1 + 10) = 1.73\,\text{bits}$

---

5. Calculate Entropy of Channels

1. **Q:**
   Given $P(Y = 1|X = 1) = 0.9$ and $P(Y = 0|X = 0) = 0.8$, find the conditional entropy $H(Y|X)$.
   A:
   $H(Y|X) = 0.5 \left(-0.9 \log_2 0.9 - 0.1 \log_2 0.1\right) + 0.5 \left(-0.8 \log_2 0.8 - 0.2 \log_2 0.2\right)$

---

6. Index Coding

1. **Q:**
   For messages W_1, W_2, W_3, find the optimal index code if:

- Client 1 knows W_2,
- Client 2 knows W_3,
- Client 3 knows W_1.
  A:
  Coded transmissions: $W_1 \oplus W_2, W_2 \oplus W_3, W_3 \oplus W_1$.

---

7. Network Coding

1. **Q:**
   In a butterfly network, compute the coded message if $A = 1$ and $B = 0$.
   A:
   Transmit $X = A \oplus B = 1$. Both sinks decode:
   $A = X \oplus B = 1, \quad B = X \oplus A = 0$

---

8. Coded Caching

1. **Q:**
   For $N = 4$, $K = 2$, and $M = 1$, find the communication load.
   A:
   $L = \frac{N(1-M/N)}{1+KM/N} = \frac{4(1-1/4)}{1+2(1/4)} = 2.4$

1.2.2 9. Gambling (after 10 Gains)

1. **Q:**
   Suppose you have achieved 10 consecutive gains and your current wealth is $W = 1000$. The probability of winning the next bet is $p = 0.55$ and the odds are $b = 2$. Apply the **Kelly Criterion** to determine the optimal bet size.

   A:
   The **Kelly Criterion** formula is:
   $f^* = \frac{bp-(1-p)}{b}$

   Substituting values:
   $f^* = \frac{2 \cdot 0.55 - 0.45}{2} = \frac{1.1 - 0.45}{2} = 0.325$

   The optimal bet size is 32.5% of your current wealth:
   $f^* \cdot W = 0.325 \cdot 1000 = 325$

---

10. MAC or Broadcast Channel (Optimal Schemes)

1. **Q:**
   For a MAC with $P_1 = 3$, $P_2 = 5$, and noise $N = 1$, find the individual rates.

A:
$R_1 \leq \frac{1}{2}\log_2\left(1+\frac{P_1}{N}\right) = 1, \quad R_2 \leq \frac{1}{2}\log_2\left(1+\frac{P_2}{N}\right) = 1.32$

---

11. Asymptotic Equipartition Property (AEP) – Picking a Small Subset of Numbers Problem**

Q: Given a random variable X with entropy $H(X) = 2\,\text{bits}$, there are $2^{10} = 1024$ possible sequences of length $n = 5$. You want to find a small subset of sequences such that their total probability is at least 0.99. How many sequences should you pick from the typical set?

Solution:

1. **Typical Set Definition:**
   The **typical set** $A_\epsilon^{(n)}$ contains sequences x^n whose probability is approximately: $P(x^n) \approx 2^{-nH(X)} = 2^{-5\cdot 2} = 2^{-10}$
2. **Total Number of Typical Sequences:**
   The number of sequences in the typical set is approximately: $|A_\epsilon^{(n)}| \approx 2^{nH(X)} = 2^{10} = 1024$
3. **Finding the Required Subset:**
   To achieve a cumulative probability of at least 0.99, we need the smallest number m of sequences such that: $m \cdot 2^{-10} \geq 0.99$

   Solving for m: $m \geq \frac{0.99}{2^{-10}} = 0.99 \cdot 1024 = 1013$
4. **Answer:**
   You need to pick at least **1013 sequences** from the typical set to ensure a cumulative probability of at least **0.99**.

This problem demonstrates how AEP helps determine the number of typical sequences necessary to capture most of the probability mass.

---

12. Coded MapReduce

1. **Q:**
   For 4 mappers and 3 reducers, calculate the communication reduction using coded MapReduce.
   A:
   $R = \frac{1}{r} = \frac{1}{3}$

This set of Q&As is designed to test both conceptual understanding and mathematical problem-solving skills in **Information Theory**.

1.3 Exercices 2

Here's a new **set of advanced questions and answers** covering all the topics on the list, focusing on more difficult mathematical problems.

1.3.1 1. Capacity Regions

1. **Q:**
   Consider a two-user Gaussian multiple access channel (MAC) with $P_1 = 3$, $P_2 = 5$, and noise $N = 2$. Find all valid rate pairs (R_1, R_2).

   A:
   The constraints are:
   $R_1 \leq \frac{1}{2}\log_2\left(1 + \frac{P_1}{N}\right), \quad R_2 \leq \frac{1}{2}\log_2\left(1 + \frac{P_2}{N}\right), \quad R_1 + R_2 \leq \frac{1}{2}\log_2\left(1 + \frac{P_1+P_2}{N}\right)$

   Calculating:
   $R_1 \leq \frac{1}{2}\log_2(1 + 1.5) \approx 0.58, \quad R_2 \leq \frac{1}{2}\log_2(1 + 2.5) \approx 0.92 \; R_1 + R_2 \leq \frac{1}{2}\log_2(1 + 4) \approx 1.16$

   The capacity region consists of all rate pairs that satisfy these inequalities.

1.3.2 2. Markov Chains

1. **Q:**
   A Markov chain has the following transition matrix:
   $P = \begin{bmatrix} 0.5 & 0.5 \\ 0.3 & 0.7 \end{bmatrix}$ Find the stationary distribution and the entropy rate.

   A:
   Step 1: Find the stationary distribution π. Solve $\pi P = \pi$:
   $\pi_1 = 0.5\pi_1 + 0.3\pi_2, \quad \pi_2 = 0.5\pi_1 + 0.7\pi_2, \quad \pi_1 + \pi_2 = 1$ Solving gives $\pi = (0.375, 0.625)$.

   Step 2: Calculate entropy rate:
   $H(X) = \sum_{i,j} \pi(i) P_{ij} \log_2 \frac{1}{P_{ij}}$ Substitution yields the entropy rate.

1.3.3 3. Maximization of Entropy

1. **Q:**
   A continuous random variable X has a Gaussian distribution with variance $\sigma^2 = 4$. Find its differential entropy and compare it to the maximum entropy of a uniform distribution over the interval $[-a, a]$.

   A:
   Step 1: Differential entropy of Gaussian:
   $h(X) = \frac{1}{2}\log_2(2\pi e\sigma^2) = \frac{1}{2}\log_2(2\pi e \cdot 4) \approx 3.06\,\text{bits}$

   Step 2: For a uniform distribution:
   $h(X) = \log_2(2a)$ To match the variance of the Gaussian, $a = 2\sqrt{3}$, so $h(X) = \log_2(4\sqrt{3}) \approx 3.17\,\text{bits}$.

1.3.4 4. Capacities of Different Channels

1. **Q:**
   Calculate the capacity of a binary symmetric channel (BSC) with crossover probability $p = 0.3$.

   A:
   $C = 1 - H(p), \quad H(p) = -p\log_2 p - (1-p)\log_2(1-p)$ Substituting $p = 0.3$:
   $H(0.3) = -(0.3\log_2 0.3 + 0.7\log_2 0.7) \approx 0.881$ $C = 1 - 0.881 = 0.119\,\text{bits}$

1.3.5 5. Calculate Entropy of Channels

1. **Q:**
   For a channel with transition matrix:
   $P(Y|X) = \begin{bmatrix} 0.9 & 0.1 \\ 0.2 & 0.8 \end{bmatrix}$, and input probabilities $P(X=1) = 0.6$, find the mutual information $I(X;Y)$.

   A:
   Step 1: Find $P(Y)$:
   $P(Y=1) = 0.6 \cdot 0.9 + 0.4 \cdot 0.2 = 0.62, \quad P(Y=2) = 0.6 \cdot 0.1 + 0.4 \cdot 0.8 = 0.38$

   Step 2: Calculate $H(Y)$ and $H(Y|X)$.
   $H(Y) = -(0.62\log_2 0.62 + 0.38\log_2 0.38)$ $H(Y|X) = 0.6 \cdot (-0.9\log_2 0.9 - 0.1\log_2 0.1) + 0.4 \cdot (-0.8\log_2 0.8 - 0.2\log_2 0.2)$

   Step 3:
   $I(X;Y) = H(Y) - H(Y|X)$

1.3.6 6. Index Coding

1. **Q:**
   For a system with 4 clients and 4 messages, each client knows all messages except the one they request. Find the optimal number of transmissions.

   A:
   Use **XOR-based** coding. Transmit:
   $W_1 \oplus W_2 \oplus W_3 \oplus W_4$ Only 1 transmission is required.

1.3.7 7. Network Coding

1. **Q:**
   In a butterfly network, if sources $A = 1$ and $B = 0$, compute the transmitted coded message and the values decoded at both sinks.

A:
Transmit: $X = A \oplus B = 1$.
Sinks decode:
$A = X \oplus B = 1, \quad B = X \oplus A = 0$

1.3.8 8. Coded Caching

1. **Q:**
   In a coded caching system with $N = 6$, $K = 3$, and $M = 2$, calculate the transmission load during the delivery phase.

   A:
   $L = \frac{N(1-M/N)}{1+KM/N} = \frac{6(1-2/6)}{1+3(2/6)} = 2$

1.3.9 9. Gambling (after 10 Gains)

1. **Q:**
   After making 10 gains, you want to maximize your long-term wealth by reinvesting a portion of your capital on each bet. Suppose the gain probability is $p = 0.6$ and the odds are $b = 1.8$. Calculate the **expected long-term growth rate** if you follow the optimal strategy.

   A:
   The **expected growth rate** G is given by:
   $G = p \log_2(1 + bf^*) + (1-p)\log_2(1 - f^*)$

   Step 1: Calculate f^*:
   $f^* = \frac{1.8 \cdot 0.6 - 0.4}{1.8} = \frac{1.08 - 0.4}{1.8} = 0.3778$

   Step 2: Substitute into the growth rate formula:
   $G = 0.6 \log_2(1 + 1.8 \cdot 0.3778) + 0.4 \log_2(1 - 0.3778)$

   Approximation yields:
   $G \approx 0.6 \cdot 0.77 + 0.4 \cdot (-0.59) \approx 0.322$

   The expected growth rate is approximately **0.322 bits** per bet.

1.3.10 10. MAC or Broadcast Channel (Optimal Schemes)

1. **Q:**
   For a broadcast channel with $P = 10$ and noise levels $N_1 = 1$, $N_2 = 4$, find the achievable rates.

   A:
   $R_1 \leq \frac{1}{2}\log_2(1 + 10), \quad R_2 \leq \frac{1}{2}\log_2\left(\frac{N_1}{N_2}\right) = 0.5$

1.3.11 11. EAP (Picking Dual Subset of Numbers)

1. **Q:**
   A random variable has entropy $H(X) = 2$. For $n = 10$, how many sequences are needed to capture 99% of the total probability?

   A:
   Probability per sequence $= 2^{-nH(X)} = 2^{-20}, \quad m \cdot 2^{-20} \geq 0.99$ Solving:
   $m \geq 0.99 \cdot 2^{20} \approx 1.04 \times 10^6$

1.3.12 12. Coded MapReduce

1. **Q:**
   In a system with 5 mappers and 4 reducers, calculate the communication cost reduction using coded MapReduce.

   A:
   $R = \frac{1}{r} = \frac{1}{4}$

These advanced problems provide a thorough challenge across **Information Theory** topics, requiring deep mathematical understanding and application of key concepts.

1.4 Exercices 3

1.4.1 Advanced Information Theory Q&A – Difficult Mathematical Problems

(Based on Elements of Information Theory, 2nd Edition by Cover & Thomas)

1.4.2 1. Capacity Regions

1. **Q:**
   Consider a two-user MAC where user 1 transmits with power $P_1 = 4$ and user 2 with $P_2 = 16$. The noise variance is $N = 1$. Derive the capacity region equations and find a rate pair (R_1, R_2) where $R_1 = 1$.

   A:
   The capacity region equations are:
   $R_1 \leq \frac{1}{2} \log_2 \left(1 + \frac{P_1}{N}\right), \quad R_2 \leq \frac{1}{2} \log_2 \left(1 + \frac{P_2}{N}\right), \quad R_1 + R_2 \leq \frac{1}{2} \log_2 \left(1 + \frac{P_1+P_2}{N}\right)$

   Substituting values:
   $R_1 \leq 1, \quad R_2 \leq 2, \quad R_1 + R_2 \leq 1.8$

   For $R_1 = 1$, R_2 must satisfy:
   $R_2 \leq 0.8$

1.4.3 2. Markov Chains

1. **Q:**
   For a Markov chain with the transition matrix:
   $P = \begin{bmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{bmatrix}$, calculate the **second-order entropy rate**, assuming the chain starts in the stationary distribution.

   A:
   Step 1: Find the stationary distribution π:
   $\pi_1 = 0.6\pi_1 + 0.3\pi_2, \quad \pi_1 + \pi_2 = 1 \quad \Rightarrow \quad \pi = \left(\frac{3}{7}, \frac{4}{7}\right)$

   Step 2: Calculate the second-order joint entropy:
   $H(X_1, X_2) = \sum_{i,j} \pi(i) P_{ij} \log_2 \frac{1}{P_{ij}}$

   Substituting values:
   $H(X_1, X_2) = \frac{3}{7}\left(0.6 \log_2 \frac{1}{0.6} + 0.4 \log_2 \frac{1}{0.4}\right) + \frac{4}{7}\left(0.3 \log_2 \frac{1}{0.3} + 0.7 \log_2 \frac{1}{0.7}\right)$

   Finally, calculate the **entropy rate** using:
   $H(X) = H(X_1, X_2) - H(X_1)$

1.4.4 3. Maximization of Entropy

1. **Q:**
   Prove that the entropy of a continuous random variable X is maximized when $X \sim \mathcal{N}(0, \sigma^2)$, by using the calculus of variations.

   A:
   The functional form of entropy is:
   $h(X) = -\int_{-\infty}^{\infty} f(x) \log f(x)\, dx$

   Applying the Euler-Lagrange equation with the constraint $\int_{-\infty}^{\infty} x^2 f(x)\, dx = \sigma^2$ leads to the solution:
   $f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$

1.4.5 4. Capacities of Different Channels

1. **Q:**
   Calculate the capacity of an AWGN channel with bandwidth $B = 5\,\text{MHz}$, signal power $P = 0.1\,\text{W}$, and noise power spectral density $N_0 = 10^{-8}\,\text{W/Hz}$.

   A:
   Capacity is given by:
   $C = B \log_2 \left(1 + \frac{P}{N_0 B}\right)$

   Substituting values:
   $C = 5 \times 10^6 \log_2 \left(1 + \frac{0.1}{5 \times 10^{-8}}\right) = 5 \times 10^6 \log_2(2001)$

   Approximation:
   $C \approx 5 \times 10^6 \times 10.97 = 54.85\,\text{Mbps}$

1.4.6 5. Calculate Entropy of Channels

1. **Q:**
   A channel has input X and output Y with the following joint probability table:
   $P(X,Y) = \begin{bmatrix} 0.3 & 0.2 \\ 0.1 & 0.4 \end{bmatrix}$ Calculate $I(X;Y)$.

   A:
   Step 1: Calculate $H(X)$ and $H(Y)$:
   $H(X) = -(0.5\log_2 0.5 + 0.5\log_2 0.5) = 1, \quad H(Y) = -(0.4\log_2 0.4 + 0.6\log_2 0.6)$

   Step 2: Calculate $H(X,Y)$:
   $H(X,Y) = -\sum_{i,j} P(x_i, y_j)\log_2 P(x_i, y_j)$

   Step 3:
   $I(X;Y) = H(X) + H(Y) - H(X,Y)$

1.4.7 6. Index Coding

1. **Q:**
   For 5 clients and 5 messages where each client is missing only the message they want, calculate the optimal number of coded transmissions.

   A:
   Use a single XOR-coded transmission:
   $W_1 \oplus W_2 \oplus W_3 \oplus W_4 \oplus W_5$

1.4.8 7. Network Coding

1. **Q:**
   In a network with two sources A and B, transmit $X = A \oplus B$. If $A = 1$ and $B = 1$, what is received and decoded at each sink?

   A:
   $X = 1 \oplus 1 = 0$

   Sinks receive $X = 0$ and decode:
   $A = X \oplus B = 0 \oplus 1 = 1, \quad B = X \oplus A = 0 \oplus 1 = 1$

1.4.9 8. Coded Caching

1. **Q:**
   For $N = 8$, $K = 4$, and $M = 2$, calculate the transmission load.

   A:
   $L = \frac{N(1-M/N)}{1+KM/N} = \frac{8(1-2/8)}{1+4(2/8)} = \frac{6}{2} = 3$

1.4.10 9. Gambling (after 10 Gains)

1. **Q:**
   Suppose you use a suboptimal betting strategy, placing a constant fraction $f = 0.5$ of your wealth on each bet. If the true optimal $f^* = 0.3778$, determine the relative difference in long-term growth rate between the optimal and suboptimal strategies.

   A:
   Step 1: Calculate the suboptimal growth rate:
   $G_{\text{suboptimal}} = 0.6 \log_2(1 + 0.9) + 0.4 \log_2(0.5)$ Approximation:
   $G_{\text{suboptimal}} \approx 0.6 \cdot 0.92 + 0.4 \cdot (-1) = 0.152$

   Step 2: Compare with optimal growth rate $G^* = 0.322$:
   $\Delta G = G^* - G_{\text{suboptimal}} = 0.322 - 0.152 = 0.17$

   The **relative difference** is:
   $\frac{\Delta G}{G^*} \approx \frac{0.17}{0.322} \approx 0.53\,(53\%)$

1.4.11 10. MAC or Broadcast Channel (Optimal Schemes)

1. **Q:**
   In a broadcast channel, the transmitter can send two messages M_1 and M_2 to two users with noise levels $N_1 = 1$ and $N_2 = 4$, respectively. The power constraint is $P = 10$. Find the achievable rate pair using **superposition coding**.

   A:

   - **Step 1:** Assign power allocations P_1 and P_2 with $P_1 + P_2 = P$.

   - **Step 2:** Calculate rates:
     For user 1 (stronger channel):
     $R_1 \leq \frac{1}{2} \log_2 \left(1 + \frac{P_1+P_2}{N_1}\right) = \frac{1}{2} \log_2(1 + 10)$ For user 2 (weaker channel):
     $R_2 \leq \frac{1}{2} \log_2 \left(1 + \frac{P_2}{N_2}\right)$
   - **Step 3:** Choose $P_1 = 6$, $P_2 = 4$:
     $R_1 = \frac{1}{2} \log_2(1 + 10) \approx 1.73, \quad R_2 = \frac{1}{2} \log_2(1 + 1) = 0.5$

---

2. **Q:**
   In a MAC with users transmitting powers $P_1 = 5$ and $P_2 = 15$, and noise variance $N = 1$, what is the achievable sum rate using **successive decoding**?

   A:

   - **Step 1:** Calculate individual rates:
     $R_1 \leq \frac{1}{2} \log_2 \left(1 + \frac{P_1}{N}\right) = \frac{1}{2} \log_2(6), \quad R_2 \leq \frac{1}{2} \log_2(16)$
   - **Step 2:** Calculate sum-rate:
     $R_1 + R_2 \leq \frac{1}{2} \log_2 \left(1 + \frac{P_1+P_2}{N}\right) = \frac{1}{2} \log_2(21) \approx 2.14\,\text{bits/symbol}$

1.4.12 11. EAP (Picking Dual Subset of Numbers)

1. **Q:**
   A random variable has entropy $H(X) = 1.5$. For $n = 20$, how many sequences are needed to cover 95% of the probability?

   A:
   Probability of each typical sequence:
   $P(x^n) = 2^{-nH(X)} = 2^{-30}$

   Solve:
   $m \cdot 2^{-30} \geq 0.95 \quad \Rightarrow \quad m \geq 0.95 \cdot 2^{30} \approx 1.02 \times 10^9$

1.4.13 12. Coded MapReduce

1. **Q:**
   In a coded MapReduce setup, there are 6 mappers and 3 reducers. Each mapper generates intermediate data needed by all reducers. How many transmissions are required without and with coding?

   A:

   - **Without coding:** Each mapper sends all data to each reducer. Total transmissions:
     $6\,\text{mappers} \times 3\,\text{reducers} = 18$
   - **With coding:** Use a coded transmission strategy, where each mapper encodes data and sends only once. Total transmissions:
     $\frac{1}{r} \cdot \text{number of intermediate blocks} = \frac{1}{3} \times 6 = 2\,\text{transmissions per reducer}$
   - Total transmissions with coding:
     6.

---

2. **Q:**
   If each reducer needs access to 3 pieces of data and each mapper can encode 2 pieces of data together, find the minimum number of coded transmissions required.

   A:

   - Data pieces per reducer = 3.

   - Each mapper can combine 2 pieces, reducing the number of transmissions:
     $\lceil 3/2 \rceil = 2\,\text{coded transmissions per reducer}$

   Total transmissions across all reducers:
   $3\,\text{reducers} \times 2 = 6$

```
[ ]:
```