Homework 1

Collaborators:

1. **Upper-bound on Entropy.** (20 points) Let $\Omega = \{1, 2, \ldots, N\}$. Suppose $X$ is a random variable over the sample space $\Omega$. For shorthand, let $p_i = \Pr[X = i]$, for each $i \in \Omega$. The random variable $X$’s entropy is defined as the following function.

   $$
   H(X) := \sum_{i \in \Omega} -p_i \cdot \ln p_i.
   $$

   Use Jensen’s inequality on the function $f(t) = \ln t$ to prove the following inequality.

   $$
   H(X) \leq \ln N.
   $$

   Furthermore, equality holds if and only if $X$ is the uniform distribution over $\Omega$.

   **Solution.**
2. **Log-sum Inequality.** (27=22+5 points)

(a) Let \( \{a_1, \ldots, a_N\} \) and \( \{b_1, \ldots, b_N\} \) be two sets of positive real numbers. Use Jensen’s inequality to prove the following inequality.

\[
\sum_{i=1}^{N} a_i \ln \frac{a_i}{b_i} \geq A \ln \frac{A}{B},
\]

where \( A := \sum_{i=1}^{N} a_i \) and \( B := \sum_{i=1}^{N} b_i \). Furthermore, equality holds if and only if \( a_i/b_i \) is identical for all \( i \in \{1, \ldots, N\} \).

**Solution.**
(b) Let $\mathcal{X}$ be a finite set and $P : \mathcal{X} \to [0,1]$ and $Q : \mathcal{X} \to [0,1]$ be two probability distributions on $\mathcal{X}$ such that for any $x \in \mathcal{X}$, $Q(x) \neq 0$. The relative entropy from $Q$ to $P$ is defined as follows:

$$D_{KL}(P \parallel Q) := \sum_{x \in \mathcal{X}} P(x) \log \frac{P(x)}{Q(x)}.$$ 

Show that for any $P$ and $Q$, it holds that $D_{KL}(P \parallel Q) \geq 0$. Moreover, state when $D_{KL}(P \parallel Q) = 0$.

Solution.
3. **Approximating Square-root.** (20 points) Our objective is to find a (meaningful and tight) lower bound for the function \( f(x) = (1 - x)^{-1/2} \) when \( x \in [0, 1) \) using a quadratic function of the form

\[
g(x) = 1 + \alpha x + \beta x^2.
\]

Use the Lagrange form of Taylor’s remainder theorem on \( f(x) \) around \( x = 0 \) to obtain the function \( g(x) \).

**Solution.**
4. **Lower-bounding Logarithm Function.** (20 points) By Taylor’s Theorem, we have seen that the following upper bound is true.

For all \( \varepsilon \in [0, 1) \) and integer \( k \geq 1 \), we have

\[
\ln(1 - \varepsilon) \leq -\varepsilon - \frac{\varepsilon^2}{2} - \cdots - \frac{\varepsilon^k}{k}
\]

We want a tight lower bound for \( \ln(1 - \varepsilon) \). Prove the following lower-bound.

For all \( \varepsilon \in [0, 1/2] \) and integer \( k \geq 1 \), we have

\[
\ln(1 - \varepsilon) \geq \left( -\varepsilon - \frac{\varepsilon^2}{2} - \cdots - \frac{\varepsilon^k}{k} \right) + \frac{\varepsilon^k}{k}
\]

(For visualization of this bound, follow this link)

**Solution.**
5. **Using Stirling Approximation.** (23 points) Suppose we have a coin that outputs heads with probability $p$ and outputs tails with probability $q = 1 - p$. We toss this coin (independently) $n$ times and record each outcome. Let $H$ be the random variable representing the number of heads in this experiment. Note that the following expression gives the probability that we get a total of $k$ heads.

$$
P[H = k] = \binom{n}{k} p^k q^{n-k}
$$

We will prove upper and lower bounds for this problem, assuming $k \geq pn$. Define $p' := k/n = (p + \varepsilon)$.

Let $P$ and $P'$ be two probability distributions on the set $\mathcal{X} = \{\text{tails, heads}\}$ such that $P(P = \text{heads}) = p$ and $P(P' = \text{heads}) = p'$.

Using the (Robbin’s form of) Stirling approximation in the lecture notes, prove the following bound.

$$
\frac{1}{\sqrt{8np' (1 - p')}} \exp \left( -nD_{KL}(P' \parallel P) \right) \leq P[H = k] \leq \frac{1}{\sqrt{2\pi np' (1 - p')}} \exp \left( -nD_{KL}(P' \parallel P) \right),
$$

where $D_{KL}(P' \parallel P)$ is the relative entropy from $P$ to $P'$ defined in question 3.

**Solution.**
6. **Computing a limit.** (20 points) Compute the following limit

\[
\lim_{n \to \infty} \sum_{j=1}^{n} \frac{\sqrt{4n^2 - j^2}}{n^2}.
\]

**Solution.**
7. **Birthday Bound.** (20 points) Intuitively, we want to claim that the following two expressions are “good approximations” of each other.

\[ f_n(t) := \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{t-1}{n}\right) \]

And

\[ g_n(t) := \exp\left(-\frac{t^2}{2n}\right) \]

To formalize this intuition, write the mathematical theorems (and then prove them) when \( t = o\left(n^{2/3}\right) \).

Hint: You may find the following inequalities helpful.

(a) \( \ln(1-x) \leq -x \), for \( x \in [0, 1) \), and

(b) \( \ln(1-x) \geq -x - x^2 \), for \( x \in [0, 1/2] \) (you already prove this identity earlier).
8. **Tight Estimation: Central Binomial Coefficient.** (Extra credit: 15 points) We will learn a new powerful technique to prove tight inequalities. As a representative example, we will estimate the central binomial coefficient. For positive integer $n$, we will prove that

$$L_n \leq \binom{2n}{n} \leq U_n,$$

where

$$L_n := \frac{4^n}{\sqrt{\pi \left(n + \frac{1}{2} + \frac{1}{4n}\right)}} \quad \quad \quad U_n := \frac{4^n}{\sqrt{\pi \left(n + \frac{1}{2} + \frac{1}{46n}\right)}}.$$

To prove these bounds, we will use the following general strategy.

(a) Define the following two sequences

\[\left\{ a_n := \binom{2n}{n}/U_n \right\}_n \quad \quad \quad \left\{ b_n := \binom{2n}{n}/L_n \right\}_n\]

(b) Prove the following limit.

\[\lim_{n \to \infty} a_n = \lim_{n \to \infty} b_n = \lim_{n \to \infty} \frac{\binom{2n}{n}}{4^n/\sqrt{\pi n}} = 1,\]

using the Stirling approximation $n! \sim \sqrt{2\pi n} \cdot (n/e)^n$.

(c) Prove $\{a_n\}_n$ is an increasing sequence.

(d) From (b) and (c), conclude that $a_n \leq 1$, implying $\binom{2n}{n} \leq U_n$.

(e) Prove $\{b_n\}_n$ is a decreasing sequence.

(f) From (b) and (e), conclude that $b_n \geq 1$, implying $\binom{2n}{n} \geq L_n$.

**Remark:** What did we achieve from this exercise? We started from the asymptotic estimate $\binom{2n}{n} \sim 4^n/\sqrt{\pi n}$. From this asymptotic estimate, we obtained explicit upper and lower bounds. We learned a powerful general technique to translate asymptotic estimates into explicit upper and lower bounds automatically.

**Solution.**