Homework 1

1. **Upper-bound on Entropy.** (20 points) Let $\Omega = \{1, 2, \dots, N\}$. Suppose $\mathbb{X}$ is a random variable over the sample space Ω. For shorthand, let us use $p_i = \mathbb{P}[\mathbb{X} = i]$, for each $i \in \Omega$. The entropy of the random variable $\mathbb{X}$ is defined to be the following function.

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

Use Jensen's inequality on the function $f(x) = \ln x$ to prove the following inequality.

$$H(\mathbb{X}) \leqslant \ln N$$

Furthermore, equality holds if and only if we have $p_1 = p_2 = \cdots = p_N$.

Solution.

2. **Log-sum Inequality.** (20 points) Let $\{a_1, \dots, a_N\}$ and $\{b_1, \dots, b_N\}$ be two sets of positive real numbers. Use Jensen's inequality to prove the following inequality

$$\sum_{i=1}^{N} a_i \ln \frac{a_i}{b_i} \geqslant A \ln \frac{A}{B},$$

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 $i \in \{1, \dots, N\}$.

Solution.

3. **Approximating Square-root.** (20 points) Our objective is to find a (meaningful and tight) upper-bound for $f(x) = (1-x)^{1/2}$ using a quadratic function of the form

$$g(x) = 1 - \alpha x - \beta x^2$$

Use the Lagrange form of the 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 $k \geqslant 1$, we have

$$\ln(1-\varepsilon) \leqslant -\varepsilon - \frac{\varepsilon^2}{2} - \cdots - \frac{\varepsilon^k}{k}$$

We are interested in obtain a tight lower-bound for $\ln(1-\varepsilon)$. Prove the following lower-bound.

For all $\varepsilon \in [0,1/2]$ and integer $k \geqslant 1$, we have

$$\ln(1-\varepsilon) \geqslant \left(-\varepsilon - \frac{\varepsilon^2}{2} - \cdots - \frac{\varepsilon^k}{k}\right) - \frac{\varepsilon^k}{k}$$

(For visualization of this bound, follow this link)

Solution.

5. **Using Stirling Approximation.** (20 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 $\mathbb{H}$ be the random variable representing the number of heads in this experiment. Note that the probability that we get k heads is given by the following expression.

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

Assume that $k > pn$, and we shall represent $p' := k/n = (p + \varepsilon)$.

Using the Stirling approximation in the lecture notes, prove the following bound.

$$\frac{1}{\sqrt{8np'(1-p')}} \exp\left(-n\mathrm{D}_{\mathrm{KL}}\left(p', p\right)\right) \leqslant \mathbb{P}\left[\mathbb{H} = k\right] \leqslant \frac{1}{\sqrt{2\pi np'(1-p')}} \exp\left(-n\mathrm{D}_{\mathrm{KL}}\left(p', p\right)\right),$$

where $\mathrm{D}_{\mathrm{KL}}\left(a, b\right)$ (referred to as the Kullback–Leibler divergence) is defined as

$$\mathrm{D}_{\mathrm{KL}}\left(a, b\right) := a \ln \frac{a}{b} + (1-a) \ln \frac{1-a}{1-b}$$

Solution.

Collaborators :