作业一

杨国威 222081202016

August 2023

task 1 Compute entropy of geometric distribution:

$$P(X = k) = p(1-p)^{k-1} \tag{1}$$

The entropy of a geometric distribution with parameter p is given by the formula:

$$H(X) = -\sum_{k=1}^{\infty} P(X = k) \log_2 P(X = k) \tag{2}$$

Substituting the probability mass function of the geometric distribution, we get:

$$H(X) = -\sum_{k=1}^{\infty} p(1-p)^{k-1} \log_2 [p(1-p)^{k-1}] \tag{3}$$

Using the properties of logarithms, we can simplify this expression as follows:

$$H(X) = -\sum_{k=1}^{\infty} p(1-p)^{k-1} [\log_2 p + (k-1) \log_2 (1-p)] \tag{4}$$

$$H(X) = -p \log_2 p \sum_{k=1}^{\infty} (1-p)^{k-1} - p \log_2 (1-p) \sum_{k=1}^{\infty} (k-1)(1-p)^{k-1} \tag{5}$$

The first summation is a geometric series with common ratio $(1-p)$, which converges to $\frac{1}{p}$. The second summation is the derivative of the first summation with respect to $(1-p)$, which converges to $\frac{1}{p(1-p)}$. Therefore, we can write:

$$H(X) = -p \log_2 p \frac{1}{p} - p \log_2 (1-p) \frac{1}{p(1-p)} \tag{6}$$

Simplifying further, we get the final answer:

$$H(X) = -\log_2 p - (1-p) \log_2 (1-p) \tag{7}$$

task 2 Construct a discrete distribution with infinite entropy

One possible way to construct a discrete distribution with infinite entropy is to use a **Zipf distribution**The general formula for a Zipf distribution is:

$$P(X = k) = \frac{1}{k^s} \frac{1}{\delta(s)} \tag{8}$$

where k is the rank of the event, s is a positive parameter that controls the shape of the distribution, and $\delta(s)$ is the Riemann zeta function, which is a normalization constant that ensures the probabilities sum to 1. The entropy of a Zipf distribution is given by:

$$H(X) = \log_2 \delta(s) + s \frac{\delta^{'}(s)}{\delta(s)} \tag{9}$$

where $\delta^{'}(s)$ is the derivative of the Riemann zeta function. As s approaches 1 from above, the entropy of the Zipf distribution diverges to infinity. This is because the distribution becomes more and more uniform as s decreases, and a uniform distribution has maximum entropy.

task 3 Compute KL divergence between Gaussians: $N(\mu_1, \sigma_1), N(\mu_2, \sigma_2)$.