Quiz 1 - Lecture 12 (Prof. Shinoda)

1. Prove the second formula in Slide 25.
2. Prove the second formula in Slide 36.
3. We think about an 1-dimension with mean μ, variance σ^2. Prove that the distribution which maximize the entropy is a Gaussian distribution.

Collaborators: None.

Exercise 1-1. Prove the second formula in Slide 25

Consider the expectations of the variation with respect to the data set values, which comes from a Gaussian distribution with parameter μ and σ^2. Prove the following formula:

$$\mathbb{E}[\sigma^2_{\text{ML}}] = \left(\frac{N-1}{N}\right)\sigma^2$$

Solution:

$$\begin{aligned}
\mathbb{E}[\sigma^2_{\text{ML}}] &= \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}(x_n - \mu_{\text{ML}})^2\right] = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\left(x_n - \frac{1}{N}\sum_{n=1}^{N}x_n\right)^2\right] \\
&= \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[\left(x_n - \frac{1}{N}\sum_{n=1}^{N}x_n\right)^2\right] \\
&= \frac{1}{N}\sum_{n=1}^{N}\mathbb{E}\left[x_n^2 - \frac{2}{N}x_n\sum_{m=1}^{N}x_m + \frac{1}{N^2}\sum_{m=1}^{N}\sum_{k=1}^{N}x_m x_k\right] \\
&= \frac{1}{N}\sum_{n=1}^{N}\left(\mathbb{E}\left[x_n^2\right] - \frac{2}{N}\mathbb{E}\left[x_n\sum_{m=1}^{N}x_m\right] + \frac{1}{N^2}\mathbb{E}\left[\sum_{m=1}^{N}\sum_{k=1}^{N}x_m x_k\right]\right)
\end{aligned} \tag{1}$$

We have these following equalities:

$$\begin{aligned}
&\mathbb{E}[x_n^2] = \mu^2 + \sigma^2 \\
&\mathbb{E}\left[x_n\sum_{m=1}^{N}x_m\right] = \mathbb{E}\left[x_n^2\right] + \sum_{m=1}^{N-1}\mathbb{E}[x_n]\mathbb{E}[x_m] = N\mu^2 + \sigma^2 \\
&\mathbb{E}\left[\sum_{m=1}^{N}\sum_{k=1}^{N}x_m x_k\right] = N\mathbb{E}[x_n^2] + 2\sum_{m=1}^{N-1}\sum_{k=m+1}^{N}\mathbb{E}[x_m]\mathbb{E}[x_k] = N^2\mu^2 + N\sigma^2
\end{aligned} \tag{2}$$

Replace results from (2) to equation (1), we have:

$$\begin{aligned}\mathbb{E}[\sigma^2_{\mathrm{ML}}] &= \frac{1}{N}\sum_{n=1}^{N}\left(\mu^2+\sigma^2-2\left(\mu^2+\frac{1}{N}\sigma^2\right)+\mu^2+\frac{1}{N}\sigma^2\right) \\ &= \left(\frac{N-1}{N}\right)\sigma^2 \ \blacksquare\end{aligned} \tag{3}$$

Exercise 1-2. Prove the second formula in Slide 36

Entropy is maximized when: $\forall i : p_i = \frac{1}{M}$

Solution: The entropy of an M-state discrete variable x can be written as:

$$\mathrm{H}(x) = -\sum_{i=1}^{M} p(x_i)\ln p(x_i) = \sum_{i=1}^{M} p(x_i)\ln\frac{1}{p(x_i)}$$

Since $\ln(x)$ is concave, therefore we can apply the reverse Jensen's inequality:

$$f\left(\sum_{i=1}^{M}\lambda_i x_i\right) \geq \sum_{i=1}^{M}\lambda_i f(x_i)$$

From the above inequality, we have:

$$\mathrm{H}(x) \leq \ln\left(\sum_{i=1}^{M} p(x_i)\frac{1}{p(x_i)}\right) = \ln M$$

The equal sign can be achieved with:

$$p_i = \frac{1}{M}$$

Hence, we have proved that the entropy is bounded by $\ln M$ and it happens when $p_i = \frac{1}{M}$.

Exercise 1-3. Prove entropy maxima for a distribution

Prove that entropy of a Gaussian distribution with mean μ and variance σ^2 is the upper bound for entrpy of some distribution with same mean and variance.

Solution: Consider a distribution $p(x)$ with fixed mean μ and variance σ^2. The entropy of the normal distribution with same entropy and variance is:

$$\mathrm{H}(\mathcal{N}(x;\mu,\sigma^2)) = \frac{1}{2}\ln(2\pi e\sigma^2)$$

Use the upper bound above, we have:

$$\begin{aligned}
\mathrm{H}(p) &\leq -\int p(x)\ln\mathcal{N}(x;\mu,\sigma^2)dx \\
&\leq -\int p(x)\ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\exp(-\frac{(x-\mu)^2}{2\sigma^2})\right) \\
&\leq \frac{1}{2\sigma^2}\int p(x)(x-\mu)^2dx + \frac{1}{2}\ln(2\pi\sigma^2) \\
&\leq \frac{1}{2}\ln(2\pi e\sigma^2) = \mathrm{H}(\mathcal{N}(x;\mu,\sigma^2))
\end{aligned}$$

Therefore, the entropy upper bound a any distribution $p(x)$ with fixed μ and σ^2 is the entropy of the Gaussian distribution with same μ and σ^2.