Homework 3

Mathematical basics for Data Science with Practice

Due Date : 10:00 AM, 09 Jun 2020

1. Assume that the Binomial distribution for x is given by

$$\mathrm{Bin}(\mathrm{m}|\mathrm{N},\mu)=\binom{\mathrm{N}}{\mathrm{m}}\mu^{\mathrm{m}}(1-\mu)^{\mathrm{N}-\mathrm{m}}$$

Show that $\mathrm{E}[\mathrm{m}]=\mathrm{N}\mu$, $\mathrm{var}[\mathrm{m}]=\mathrm{N}\mu(1-\mu)$. **(10pts)**

2. In the information theory, the **entropy** of the random variable x is defined by

$$\mathrm{H}[x]=-\sum_{x}p(x)\log p(x)$$

where $p(x)$ is the probability of x. **(10pts)**

(a) Explain the meaning of the entropy $\mathrm{H}[x]$.

(b) Show that the entropy $\mathrm{H}[x]$ of a Bernoulli distributed random binary variable x is given by

$$\mathrm{H}[x]=-\mu\log\mu-(1-\mu)\log(1-\mu)$$

where $p(x|\mu)=\mu^{x}(1-\mu)^{1-x}$.

3. Note that the Bernoulli distribution given by $\mathrm{Bern}(x|\mu)=\mu^{x}(1-\mu)^{1-x}$ is not symmetric between the two variables. For endowing the symmetric property, the following equivalent formulation for which $x\in\{-1,1\}$

$$p(x|\mu)=\left(\frac{1-\mu}{2}\right)^{(1-x)/2}\left(\frac{1+\mu}{2}\right)^{(1+x)/2}$$

is also used, where $\mu\in[-1,1]$. Show that the followings. **(10pts)**

(a) $\sum_{x}p(x|\mu)=1$

(b) Compute the mean $\mathrm{E}[x]$, $\mathrm{var}[x]$.

4. **(Preliminary for Gaussian distribution)** Show that

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}\,.$$

Hint : Double integral & Change of variable with polar coordinate. **(10pts)**

5. (Python Example) Let us consider the function given by

$$g(w_1, w_2) = w_1^2 + w_2^2 + 2\sin^2(1.5(w_1 + w_2)) + 2$$

Write a python code for finding minimum of $g(w_1, w_2)$ starting at $\boldsymbol{w}_0 = (3,3)$ with the Gradient descent method. The code **MUST** include the followings.

(a) Stepsize $t = 10^{-2}$, 10^{-1}, 10^{0} (Comment1)

(b) Backtracking (Comment2)

Compare the results from (a) and (b). **(10pts)**