Elliot Smith

NML 502

Homework 2

Problem 1

a. $f(x) = \dfrac{1}{1+e^{-ax}}$, $\dfrac{df(x)}{dx} = af(x)[1-f(x)]$

$f(x) = \dfrac{g(x)}{h(x)}$, $f'(x) = \dfrac{g'(x)h(x) - g(x)h'(x)}{[h(x)]^2}$

$$f'(x) = \frac{-(-ae^{-ax})}{[1+e^{-ax}]^2} = \frac{ae^{-ax}}{[1+e^{-ax}]^2} = \frac{a}{(1+e^{-ax})} \cdot \frac{e^{-ax}}{(1+e^{-ax})}$$

$$= \frac{a}{(1+e^{-ax})} \cdot \frac{(1+e^{-ax})-1}{(1+e^{-ax})} = \frac{a}{(1+e^{-ax})}\left[\frac{(1+e^{-ax})}{(1+e^{-ax})} - \frac{1}{(1+e^{-ax})}\right]$$

$$= \frac{a}{(1+e^{-ax})}\left[1 - \frac{1}{(1+e^{-ax})}\right] = af(x)[1-f(x)]$$

b. $f(z) = \dfrac{e^{bz}-e^{-bz}}{e^{bz}+e^{-bz}}$

$$f'(z) = \frac{(be^{bz}+be^{-bz})(e^{bz}+e^{-bz}) - (e^{bz}-e^{-bz})(be^{bz}-be^{-bz})}{[e^{bz}+e^{-bz}]^2}$$

$$= \frac{b(e^{bz}+e^{-bz})(e^{bz}+e^{-bz}) - b(e^{bz}-e^{-bz})(e^{bz}-e^{-bz})}{[e^{bz}+e^{-bz}]^2}$$

$$= b\left[\frac{(e^{bz}+e^{-bz})^2 - (e^{bz}-e^{-bz})^2}{[e^{bz}+e^{-bz}]^2}\right]$$

$$= b\left[1 - \frac{(e^{bz}-e^{-bz})^2}{(e^{bz}+e^{-bz})^2}\right]$$

$$= b\left[1 - \left(\frac{e^{bz}-e^{-bz}}{e^{bz}+e^{-bz}}\right)^2\right]$$

$$= b[1-(f(z))^2]$$

c. To achieve this, we would need to scale the inputs by $\frac{1}{a}$.

Problem 2

i. 1.8

ii. 1

iii. 0.8581

Problem 3

a. Yes

b. No

c. No

d. Yes

Problem 5

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

entropy $\rightarrow E[-\log p(x)]$.

$$= \int_{-\infty}^{\infty} p(x)(-\log p(x))\, dx \longrightarrow (-\log p(x)) = \frac{1}{2}\log(2\pi\sigma^2) + \frac{1}{2\sigma^2}(x-\mu)^2$$

$$= \frac{1}{2}\log(2\pi\sigma^2)\int_{-\infty}^{\infty} p(x)\, dx + \frac{1}{2\sigma^2}\int_{-\infty}^{\infty}(x-\mu)^2 dx \rightarrow \int_{-\infty}^{\infty} p(x)\, dx = 1$$

$$\rightarrow \int_{-\infty}^{\infty}(x-\mu)^2\, dx = \sigma^2$$

$$= \frac{1}{2}\log(2\pi\sigma^2) + \frac{1}{2}$$

$$= \frac{1}{2}(1 + \log(2\pi\sigma^2))$$

Problem 6

a. $EX = \int_a^b x \frac{1}{b-a}\,dx$

$= \frac{1}{b-a}\int_a^b x\,dx$

$= \frac{1}{b-a}\left[\frac{1}{2}x^2 \Big|_a^b\right]$

$= \frac{1}{2(b-a)}\left[b^2-a^2\right]$

$= \frac{1}{2(b-a)}(b-a)(b+a)$

$= \frac{b+a}{2}$

$EX^2 = \int_a^b x^2 \frac{1}{b-a}\,dx$

$= \frac{1}{b-a}\int_a^b x^2\,dx$

$= \frac{1}{b-a}\left[\frac{1}{3}x^3\Big|_a^b\right]$

$= \frac{1}{3(b-a)}\left[b^3-a^3\right]$

$= \frac{1}{3(b-a)}(b-a)\left[b^2+ab+a^2\right]$

$= \frac{b^2+ab+a^2}{3}$

$$\operatorname{Var}X = \frac{b^2+ab+a^2}{3} - \frac{b^2+2ab+a^2}{4} = \frac{4b^2+4ab+4a^2-3b^2-6ab-3a^2}{12} = \frac{b^2-2ab+a^2}{12} = \frac{(b-a)^2}{12}$$

Entropy $= \int_a^b \frac{1}{b-a}\log(b-a)\,dx$

$= \frac{1}{b-a}\log(b-a)\int_a^b 1\,dx$

$= \frac{1}{b-a}\log(b-a)(b-a)$

$= \log(b-a)$

b. Unif $\rightarrow h(x) = \log(b-a)$

Normal $\rightarrow h(x) = \frac{1}{2}(1+\log(2\pi\sigma^2))$

Variance $=1$

Unif $\rightarrow h(x) = \log(\sqrt{12}-0) = 1.2425$

Normal $\rightarrow h(x) = \frac{1}{2}(1+\log(2\pi)) = 1.4189$

Variance $= 100$

Unif $\rightarrow h(x) = \log(\sqrt{1200}-0) = 3.5458$

Normal $\rightarrow h(x) = \frac{1}{2}(1+\log(2\pi\cdot 100)) = 3.7215$

Variance $= 10{,}000$

Unif $\rightarrow h(x) = \log(\sqrt{120{,}000}-0) = 5.8476$

Normal $\rightarrow h(x) = \frac{1}{2}(1+\log(2\pi\cdot 10{,}000)) = 6.0241$

So, the differential entropy of Gaussian distribution is larger!

Problem 7

Consider two continuous pdfs on an interval I in real numbers, such that $p \geq 0$ and $q > 0$, so:

$$-\int_I p \log p \, dx \leq -\int_I p \log q \, dx,$$

where equality holds if and only if $p(x) = q(x)$ for all x

Let p be any pdf $(x_1, \ldots, x_n)$ and $p_i = p(x_i)$, when we let $q_i = \frac{1}{n}$:

$$-\sum_{i=1}^{n} p_i \log q_i = \sum_{i=1}^{n} p_i \log n = \log n \rightarrow \text{the entropy of } q$$

Therefore, $h(p) \leq h(q)$ is equal if and only if p is uniform and helps us to conclude that the Uniform pdf maximizes the differential entropy of the random variable.

Problem 8

a. Bivariate Normal

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[\frac{-1}{2(1-\rho^2)}\left(\left[\frac{x-\mu_x}{\sigma_x}\right]^2 + \left[\frac{y-\mu_y}{\sigma_y}\right]^2 - 2\rho\left[\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y}\right]\right)\right]$$

In the case they are uncorrelated $(\rho = 0)$, it is clear to see that it resolves to a product of two normal densities. If we know that the joint distribution can be written as a product of non-negative functions, we know that x and y are independent.

b. $X \perp Y$ $\qquad p(x,y) = p(x)p(y)$

$h(x,y) = -\iint p(x,y) \log p(x,y) \, dx dy$

$h(x) = -\int p(x) \log p(x) \, dx$

$h(y) = -\int p(y) \log p(y) \, dy$

$h(x,y) = -\iint p(x,y) \log p(x) p(y) \, dx dy$

$= -\iint p(x,y) \log p(x) \, dx dy - \iint p(x,y) \log p(y) \, dx dy$

$= -\iint p(x,y) \, dy \log p(x) \, dx - \iint p(x,y) \, dx \log p(y) \, dy$

So: $h(x,y) = h(x) + h(y)$