Seokwun Kim   1.31, 1.32, 1.33, 1.35, 1.39, 1.40, 1.41

ex 1.31) two variables x and y having joint distribution $p(x,y)$
differential entropy the pair of variable satisfies

$$H[x,y] \leq H[x] + H[y]$$ iff x and y are statistically independent.

sol) $I(x;y) = H(y) - H(y|x)$ ↳ Conditional and joint entropy

↓

mutual information $I(x;y) \geq 0$

└ Jensen's Inequality

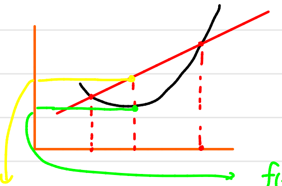
$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2)$$

$I(X;Y) = H(X) - H(X|Y)$

$= H(Y) - H(Y|X)$

$= H(X) + H(Y) - H(X,Y)$

$= H(X,Y) - H(X|Y) - H(Y|X)$



$f(tx_1 + (1-t)x_2)$

$tf(x_1) + (1-t)f(x_2)$

x and y independent → $p(x,y) = p(x)p(y)$

$$I(X;Y) = H(X,Y) = \iint p(x,y) \ln p(x,y)\, dx\, dy$$

$$I(X;Y) = \sum_{x,y} P_{xy}(x,y) \log \frac{P_{xy}(x,y)}{P_x(x) P_y(y)}$$

$$= \iint p(x)p(y) \{\ln p(x) + \ln p(y)\}\, dx\, dy$$

$$= H(X) + H(Y)$$

$$H(x|y) = \sum_y P_y(y) \left[- \sum_x P_{x|y}(x|y) \log (P_{x|y}(x|y)) \right]$$

$$\therefore I(x;y) = H(X) - H(X|Y)$$

$$\begin{aligned}
\mathrm{I}(X;Y) &= \sum_{x\in\mathcal{X},y\in\mathcal{Y}} p_{(X,Y)}(x,y)\log\frac{p_{(X,Y)}(x,y)}{p_X(x)p_Y(y)} \\
&= \sum_{x\in\mathcal{X},y\in\mathcal{Y}} p_{(X,Y)}(x,y)\log\frac{p_{(X,Y)}(x,y)}{p_X(x)} - \sum_{x\in\mathcal{X},y\in\mathcal{Y}} p_{(X,Y)}(x,y)\log p_Y(y) \\
&= \sum_{x\in\mathcal{X},y\in\mathcal{Y}} p_X(x)p_{Y|X=x}(y)\log p_{Y|X=x}(y) - \sum_{x\in\mathcal{X},y\in\mathcal{Y}} p_{(X,Y)}(x,y)\log p_Y(y) \\
&= \sum_{x\in\mathcal{X}} p_X(x)\left(\sum_{y\in\mathcal{Y}} p_{Y|X=x}(y)\log p_{Y|X=x}(y)\right) - \sum_{y\in\mathcal{Y}}\left(\sum_{x} p_{(X,Y)}(x,y)\right)\log p_Y(y) \\
&= -\sum_{x\in\mathcal{X}} p(x)\mathrm{H}(Y \mid X=x) - \sum_{y\in\mathcal{Y}} p_Y(y)\log p_Y(y) \\
&= -\mathrm{H}(Y\mid X) + \mathrm{H}(Y) \\
&= \mathrm{H}(Y) - \mathrm{H}(Y\mid X).
\end{aligned}$$

$$H(x) = -\sum_x P_x(x) \log P_x(x)$$

ex1.32) vector x of continuous variables with distribution $p(x)$ corresponding $H[x]$

transformation of x to obtain $y = Ax$ $H[y] = H[x] + \ln|A|$

entropy

sol)

$$H[y] = -\int p(y) \ln p(y)\, dy$$

$\Delta \to 0$

nonlinear

change of variable

jacobian factor

differential entropy

$\hookrightarrow \lim_{\Delta \to 0} \{ \sum_i p(x_i) \Delta \ln f(x_i))$

$x = g(y) \to f(y) = f(g(x))$

density $p(x)$ / $p(y)$

$$p(x)\delta x = p(y)\delta y \to \boxed{p(y) = p(x)\left|\frac{dx}{dy}\right|}$$

$x + \delta x$, $y + \delta y$

$|A|^{-1} = \frac{1}{|A|}$

$$p(x) = p(y) \left|\frac{dy}{dx}\right| = p(y)|A|$$

$$H(y) = -\int p(y) \ln p(y) dy = -\int p(y) \ln [p(x)|A|^{-1}]\, dy$$

$$= H[x] + \ln|A|$$

$$H[x] = -\int p(x) \ln p(x) dx$$

Entropy Maximize : Discrete - Uniform

Continuous - Gaussian

Minimize -- zero!

random variable X and Y independent.

$h(x) = -\log p(x)$

$$H[x] = -\sum p(x) \ln p(x) = E_p[-\ln p(x)]$$

uniform, normal etc. Various kind

. property

$H(x+c) = H(x)$

$H(ax) = H(x) + \ln|a|$

$H(AX) = H(X) + \ln|\det(A)|$

Properties

X random variable (r.v) $\quad F(x) = \int_{-\infty}^{x} f_x(t)\,dt \quad y = g(x) \rightarrow x = g^{-1}(y)$

$F_Y(y) = F_x(g^{-1}(y)) \quad f_Y(y) = \frac{dF(y)}{dy} = f_x(g^{-1}(y)) \left| \frac{dg^{-1}(y)}{dy} \right| = f_x(x) \left| \frac{dx}{dy} \right|$

non-increasing under function variable let $g(X)$

$H(g(X)) \leq H(X)$

with equality iff g invertible

$H(X, g(X)) = H(X, g(X))$

$H(X) + H(g(X)|X) = H(g(X)) + H(X|g(X))$

$H(X) - H(g(X)) = H(X|g(X)) \geq 0$

ex 1.33) suppose conditional entropy $H[y|x]$ two discrete random variable x, y zero.

for all values x s.t $p(x) > 0$. $p(y|x) \neq 0$

sol) $H(y|x) = -\int f(x,y) \ln f(x|y) dx dy = -\int f(x,y) \ln \frac{f(x,y)}{f(y)} dx dy$

conditional entropy

$= H(X,Y) - H(Y)$

where, $f(x,y)$ joint density

$H(y|x) = -\int\int p(Y,x) \ln (y|x) dy dx$

$H(X,Y) = H(y|X) + H(X)$

↓

chain rule.

↓

$H(X,Y) = H(X|Y) + H(Y)$

def. $H(X,Y) = -E[\ln p(X,Y)]$

$= -\sum_x \sum_y p(X,Y) \ln P(X,Y)$

$H(Y|X) = -E[\ln P(Y|X)] = -\sum_x \sum_y p(x,y) \ln p(y|x)$

$= -\sum_x \sum_y p(X,Y) \ln P(Y|X) = -\sum_x p(x) \sum_y p(y|x) \ln p(y|x) = \sum_x p(x) \cdot H(Y|X=x)$

$E[\ln P(X,Y)] = E[\ln p(X)] + E[\ln P(Y|X)]$

equal 0,

s.t $p(x) \neq 0$ $p(y|x)$

are either 0 and 1

ex1.35) result (1.106), (1.107) show that entropy univariate Gaussian (1.109) given (1.110)

sol) $\delta x \downarrow dx$

$$H(X) = -\ln(\delta x) - \sum_i p(X=x_i)\,\delta x \,\ln(p(X=x_i))$$

$$H(X) = -\ln(dx) - \int p(X=x)\ln(p(X=x))\,dx$$

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

$$E(x) = E[-\ln p(x)]$$ pro. density f.

$$= \int_{-\infty}^{\infty} p(x)(-\ln p(x))\,dx$$

$$-\ln p(x) = \frac{1}{2}\ln(2\pi V) + \frac{1}{2V}(x-\mu)^2$$

V: variance (σ^2)

μ: mean

$$\frac{1}{2}\ln(2\pi V)\int_{-\infty}^{\infty} p(x)\,dx + \frac{1}{2V}\int_{-\infty}^{\infty}(x-\mu)^2 p(x)\,dx$$

Variance $= V$

1 : properties

$$\therefore \frac{1}{2}(\ln(2\pi V) + 1)$$

f(rear) = (1+rear) % Q_SIZE

front = (front+1) % Q

(106) ... $\int_{-\infty}^{\infty} x\,p(x)\,dx = \mu$

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

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

(110) ... $H(x) = \frac{1}{2}[1 + \ln(2\pi\sigma^2)]$

ex1.37) using defin (1.111) together with product rule prob, prove result (1.112)

sol)

$$H[Y|X] = -\iint p(X,Y)\ln p(X,Y)\,dY\,dX = -\iint p(X,Y)\ln\big(p(Y|X)\,p(X)\big)\,dY\,dX$$

$$= -\iint p(X,Y)\big(\ln p(Y|X) + \ln p(X)\big)\,dY\,dX$$

$$= -\iint p(X,Y)\ln p(Y|X)\,dX\,dY - \iint p(X,Y)\ln p(X)\,dX\,dY$$

$\downarrow$

$p(X)\ln p(X)$

$$= \boxed{H[Y|X] + H[X]}$$

$p(X,Y) = p(Y|X)\,p(X)$

ex1.40) By applying Jensen's inequality (1.115) $f(x) = \ln x$

mean of a set of real number is never less their geometric mean.

(1.115) $f\left(\sum_{i=1}^{M} \lambda_i x_i\right) \le \sum_{i=1}^{M} \lambda_i f(x_i)$ → Jensen's inequality

$f(E[x]) \le E[f(x)]$

↓

$f\left(\int x p(x) dx\right) \le \int f(x) p(x) dx$

arithemic

↓

A $\bar{x}_A = \frac{1}{K} \sum_{k}^{K} x_k$

Geometric

↓

G $\bar{x}_G = \left(\prod_{k}^{K} x_k\right)^{1/K}$

↳ $\sqrt[K]{x_1 \cdot x_2 \cdots x_k}$

nth root of the number

$\ln \bar{x}_A = \ln\left(\frac{1}{K} \sum_{k}^{K} x_k\right)$

$\ln \bar{x}_G = \ln\left(\prod_{k}^{K} x_k\right)^{1/K} = \frac{1}{K} \sum_{k}^{K} \ln x_k$

set of point $\{x_i\}$ $\lambda_i \ge 0$ $\sum_i \lambda_i = 1$.

---

24 / 4.2

96

(학기 per 4.

96 + 2