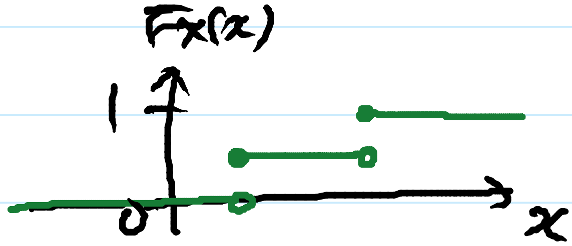
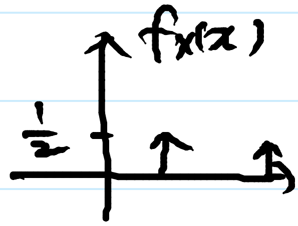
Lec. 23

2013년 5월 8일 수요일
오후 8:56

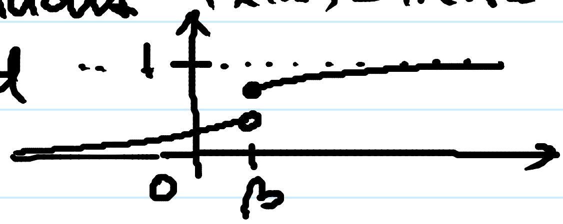
Review

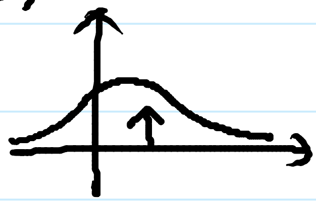




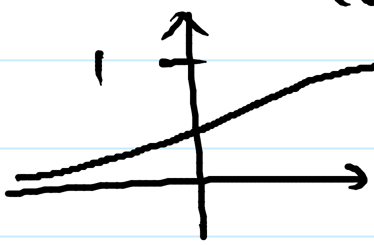
r.v. — discrete
— continuous
— mixed

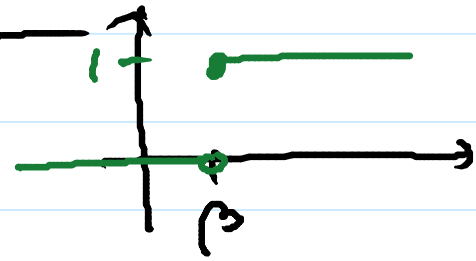
$F_X(x) \triangleq \Pr(X \le x)$





$$F_X(x) = \alpha \underbrace{F_{X_1}(x)}_{\text{continuous}} + (1-\alpha)\underbrace{F_{X_2}(x)}_{\text{discrete}}$$





- $H(X)$
  $H(X,Y)$
  $H(X|Y)$
  $I(X;Y)$, $I(X;Y|Z)$
  $D(P_X \| P_Y)$, $D(P_{XY} \| P_{ZW})$

- 

$Z \sim N(0, \sigma_Z^2)$

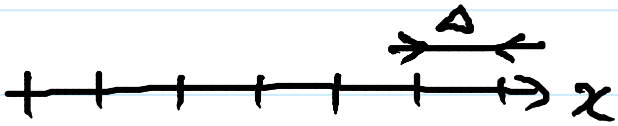
normal, Gaussian

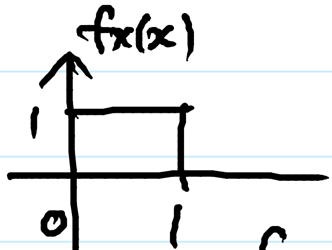
Page 01

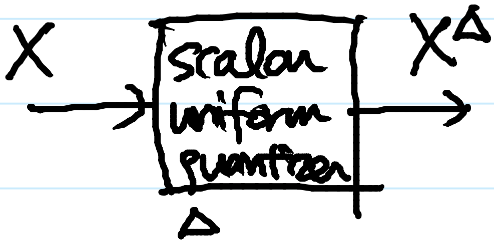
2013년 5월 8일 수요일
오후 8:56

- Given jointly distributed continuous random variables $X, Y, Z, X_1, X_2, \dots Y_1, Y_2, \dots$

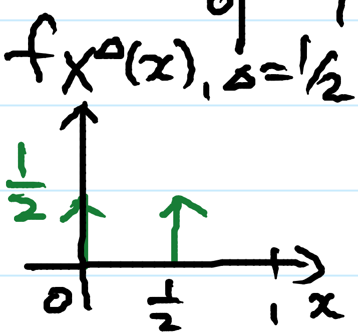
$$\cancel{H(X) \triangleq E\left\{\log \frac{1}{p_X(x)}\right\}}$$

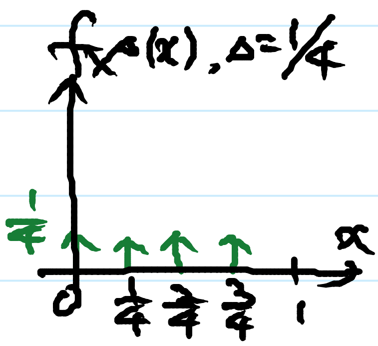






continuous
rv





$$H(X^\Delta) = \sum_i \Big(F_X(\Delta_{i+1}) - F_X(\Delta_i)\Big) \log \frac{1}{(\qquad)}$$

$$\simeq \sum_i f_X(\Delta_i)\Delta \log \frac{1}{f_X(\Delta_i)\Delta}$$

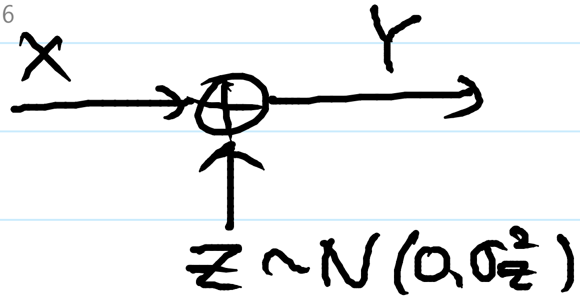
$$= \sum_i f_X(\Delta_i)\Delta \log \frac{1}{f_X(\Delta_i)}$$

$$+ \sum_i f_X(\Delta_i)\Delta \log \frac{1}{\Delta}$$

$$\simeq \int_{-\infty}^{\infty} f_X(x) \log \frac{1}{f_X(x)}\, dx$$

$$+ \log \frac{1}{\Delta}$$

$$E\left\{\log \frac{1}{f_X(x)}\right\} \triangleq h(X)$$

- 

$$I(X^{\Delta x}; Y^{\Delta y}) = H(X^{\Delta x}) + H(Y^{\Delta y}) - H(X^{\Delta x}, Y^{\Delta y})$$

$$\simeq h(X) + \log \frac{1}{\Delta x} + h(Y) + \log \frac{1}{\Delta y} - h(X,Y) - \log \frac{1}{\Delta x \Delta y} \to h(X) + h(Y) - h(X,Y)$$

$$I(X;Y) \triangleq h(X) + h(Y) - h(X,Y) = h(X) - h(X|Y) = h(Y) - h(Y|X)$$

where $h(X|Y) \triangleq E\left\{\log \frac{1}{f_{X|Y}(X|Y)}\right\}$

- $h(X) = E\left\{\log \frac{1}{f_X(X)}\right\}$ could be +, −, not defined.

$$h(X,Y) = E\left\{\log \frac{1}{f_{X,Y}(X,Y)}\right\}$$

$h(X|Y) =$

$$I(X;Y) = E\left\{\log \frac{f_{X,Y}(X,Y)}{f_X(X) f_Y(Y)}\right\} \geq 0$$

$$D(f_X \| f_Y) = E\left\{\log \frac{f_X(X)}{\cdots}\right\} \geq 0$$

$$D(f_X \| f_Y) = E\left\{\log \frac{f_X(X)}{f_Y(X)}\right\} \geq 0$$

- Properties

$$h(X_1, X_2, \cdots, X_n) = \sum_{i=1}^{n} h(X_i | X_1, X_2, \cdots, X_{i-1})$$

$$I(X;Y) = h(X) - h(X|Y) \geq 0$$

$$h(X_1, X_2, \cdots, X_n) \leq \sum_{i=1}^{n} h(X_i)$$ independence upper bound

$$h(X+c) = h(X)$$

$$h(aX) = h(X) + \log|a|$$

$$h(X_1, X_2, \cdots, X_n) = h(\underline{X}) \quad \& \quad \underline{Y} = A\underline{X}$$

$$h(A\underline{X}) = h(\underline{X}) + \log|\det A|$$

In particular, A is invertible.

$$y = ax \Rightarrow dy = a\,dx$$

$$\int_{-\infty}^{\infty} g(x)\,dx = \int_{-\infty}^{\infty} g\left(\frac{y}{a}\right) \frac{dy}{|a|}$$

1×1, diagonal

$$\underline{Y} = A\underline{X} \Rightarrow d\underline{Y} = \det A\, d\underline{X}$$

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(\underline{x})\,d\underline{x} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(A^{-1}\underline{y}) \frac{d\underline{y}}{|\det A|}$$

Jacobian determinant

2013년 5월 8일 수요일

오후 8:56

- $X \sim N(m, \sigma^2)$

$$h(X) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}} \log \frac{1}{\cdots} dx$$

$$\log \frac{1}{\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}} = \frac{1}{2}\log(2\pi\sigma^2) + \log e^{\frac{(x-m)^2}{2\sigma^2}}$$

$$= \frac{1}{2}\log(2\pi\sigma^2) + \frac{(x-m)^2}{2\sigma^2}\log e$$

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

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

- $\underline{X} \sim N(\underline{\mu}, C)$, $C > 0$

If $C = \mathrm{diag}(\sigma_1^2, \sigma_2^2, \cdots, \sigma_n^2)$, then

$$h(\underline{X}) = \frac{1}{2}\log\{(2\pi e)^n \sigma_1^2\sigma_2^2\cdots\sigma_n^2\}$$

In general

$$h(\underline{X}) = \frac{1}{2}\log\{(2\pi e)^n \det C\}$$

- Theorem

If $\underline{X}$ satisfies $E\{\underline{X}\} = \underline{\mu}$ & $\mathrm{Cov}\{\underline{X}\} = C$, then

Page 05

2013년 5월 8일 수요일

오후 8:56

$$h(\underline{X}) \le \frac{1}{2}\log\{(2\pi e)^n \det C\}$$

(proof) Let $\underline{Y} \sim N(\underline{\mu}, C)$

$$D(f_{\underline{X}} \| f_{\underline{Y}}) = E\left\{\log \frac{f_{\underline{X}}(\underline{X})}{f_{\underline{Y}}(\underline{X})}\right\}$$

$$= -E\left\{\log \frac{1}{f_{\underline{X}}(\underline{X})}\right\} + E\left\{\log \frac{1}{f_{\underline{Y}}(\underline{X})}\right\}$$

$$= -h(\underline{X}) + \int_{\underline{x}} \cancel{f_{\underline{X}}(\underline{x})}\, f_{\underline{Y}}(\underline{x}) \log \frac{1}{\frac{1}{\sqrt{2\pi}^n \sqrt{\det C}} \exp\left(-(\underline{x}-\underline{\mu})^T C^{-1} (\underline{x}-\underline{\mu})\right)} d\underline{x}$$

$$= -h(\underline{X}) + h(\underline{Y}) \ge 0$$

- The Asymptotic Equipartition Property (AEP) for a continuous random variable X

the support of a ft $f_X(x)$.

Given $f_X(x)$, define $S \triangleq \{x : f_X(x) > 0\}$

(i) The ε-typical set $A_\varepsilon^{(n)}$ is defined as

$$A_\varepsilon^{(n)} \triangleq \left\{(x_1, x_2, \cdots, x_n) \in S^n : \left| -\frac{1}{n}\log \prod_{i=1}^{n} f_X(x_i) - h(X)\right| < \varepsilon\right\}$$

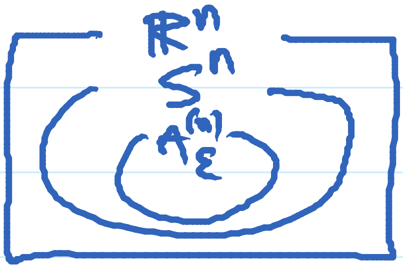
$$Y_i = -\log f_X(X_i)$$

$(Y_1+Y_2+\cdots+Y_n)/n \longrightarrow h(X)$ in prob.

Page 06

2013년 5월 8일 수요일

오후 8:56



(ii) • $\Pr(\underline{X}^{(n)} \in A_\varepsilon^{(n)}) > 1-\varepsilon$ for s.l.n.

• If $\underline{x}^{(n)} \in A_\varepsilon^{(n)}$, then

$$2^{-n(h(X)+\varepsilon)} \le f_{\underline{X}^{(n)}}(\underline{x}^{(n)}) \le 2^{-n(h(X)-\varepsilon)}$$

• $\mathrm{Vol}(A_\varepsilon^{(n)})$

$$1 = \int_{\underline{x} \in S^n} f_{\underline{X}}(\underline{x})\, d\underline{x}$$

$$\ge \int_{\underline{x} \in A_\varepsilon^{(n)}} f_{\underline{X}}(\underline{x})\, d\underline{x}$$

$$\ge \int_{\underline{x} \in A_\varepsilon^{(n)}} 2^{-n(h(X)+\varepsilon)}\, d\underline{x}$$

$$= 2^{-n(h(X)+\varepsilon)}\, \mathrm{Vol}(A_\varepsilon^{(n)})$$

$$\Rightarrow \mathrm{Vol}(A_\varepsilon^{(n)}) \le 2^{n(h(X)+\varepsilon)}$$

$$1-\varepsilon < \int_{\underline{x} \in A_\varepsilon^{(n)}} f_{\underline{X}}(\underline{x})\, d\underline{x} \quad \text{for s.l.n}$$

$$< \int_{\underline{x} \in A_\varepsilon^{(n)}} 2^{-n(h(X)-\varepsilon)}\, d\underline{x}$$

$$= \mathrm{Vol}(A_\varepsilon^{(n)})\, 2^{-n(h(X)-\varepsilon)}$$

$$\Rightarrow \mathrm{Vol}(A_\varepsilon^{(n)}) > (1-\varepsilon)\, 2^{n(h(X)-\varepsilon)} \quad \text{for s.l.n}$$