20CS6037 Machine Learning

MLE, MAP, Bayesian Reasoning - Chapter 3 & 5 (Lecture 5: 9/9/14)

Lecturer: Anca Ralescu *Scribes: Khaldoon Ashouiliy, Kyungmook Park*

Section 1: Conditional Independence

Section 2: Transformation of Random Variables

Section 3: General Transformations

Section 4: Monte Carlo Approximation

Section 5: Entropy

Section 6: Mutual Information

Section 1: Conditional Independence

X,Y r.v. X $\perp Y$: X and Y are independent

<u>Def</u>

X $\perp$ Y $\Leftrightarrow$ P(X,Y) = P(X)P(Y)

This really means $\{w \in S \mid X(w) = a\}$, $\{w \in S \mid Y(w) = b\}$
are independent event for all a $\in$ Range(x); b $\in$ Range(Y)

<u>Notation</u>

X $\perp$ Y | 2 : X and Y are ***Conditionally Independent*** **(CI)** given 2

P(X $\perp$ Y | 2) $\Leftrightarrow$ P(X, Y | 2) = P(X | 2)P(Y | 2)

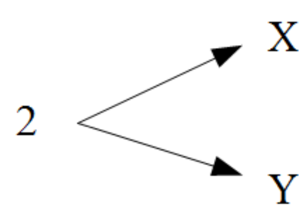

Figure 1: X and Y are CI given 2

<u>P</u>(X, Y) : Joint Distribution of X and Y
$\{w \in S \mid X(w) = a\}$, $\{w \in S \mid Y(w) = b\}$
Can extend to <u>P</u>($X_1, \ldots, X_D$) : CDF, PDF/PMF

$$\begin{aligned} \text{COV(X, Y)} &\triangleq \text{E[(X-E(X))(Y-E(Y))]} \\ &= \text{E[XY - XE(Y) - YE(X) + E(X)E(Y)]} \\ &= \text{E(XY) = E(X)E(Y)} \end{aligned}$$

For a vector X = (X_1, X_2, X_3)

$$Cov[X] = \begin{bmatrix} Var(X_1) & Cov(X_1, X_2) & Cov(X_1, X_3) \\ Cov(X_2, X_1) & Var(X_2) & Cov(X_2, X_3) \\ Cov(X_3, X_1) & Cov(X_3, X_2) & Var(X_3) \end{bmatrix}$$

Cov $\in$ (0, ∞)

<u>Correlation</u> ρ(X,Y) = $\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$ $\in$ [-1,1]

X, Y independent $\Rightarrow$ Cov(X,Y) = 0 $\Rightarrow$ ρ(X,Y) = 0 (Uncorrelated)

Independence $\rightleftharpoons$ Uncorrelated

ex)
X $\sim$ u(-1,1) ; Y = X^2 $\Rightarrow$ X, Y are dependent
ρ(X,Y) = 0

E(X) = $\frac{-1+1}{2}$ = 0 Var(X) = $\frac{(1-(-1)^2}{12}$ = $\frac{4}{12}$ = $\frac{1}{3}$

$$\begin{aligned} E(Y) \quad &= \int_{-1}^{1} x^2 f(x)dx \\ &= \int_{-1}^{1} x^2 (\tfrac{1}{2})dx = \tfrac{1}{2}\tfrac{x^4}{4}|_{-1}^{1} = \tfrac{2}{6} = \tfrac{1}{3} \end{aligned}$$

E(XY) = $\int_{-1}^{1} x^3 \frac{1}{2} dx = \frac{1}{2}\frac{x^4}{4}|_{-1}^{1} = \frac{1}{2} \cdot 0$

Cov(X,Y) = E(XY) - E(X)E(Y) = 0 - 0 $\cdot$ $\frac{1}{3}$ = 0 - 0 = 0

Section 2: Transformation of Random Variables

X $\sim$ P(X) p:pdf
Y = f(X) What is the distribution of Y?
$P_Y(Y \leq y) = P_Y(f(X) \leq y) = \underline{P}_X(X \leq f^{-1}(y)) = P(f^{-1}(y))$

ex) Y = aX+b $\Rightarrow$ f(x) = aX+b

f^{-1}(y) = $\frac{y-b}{a}$ X = -1 Y = b-a ; X = 1 Y = a+b

P$_Y$(y) = P$_X$ ($\frac{y-b}{a}$)

ex) X $\sim$ u(-1,1)

$$P_X(x) = \begin{cases} 0 & x \leq -1 \\ \frac{1}{1-(2-1)} & -1 < X < 1 \\ 0 & x \geq 1 \end{cases}$$

$$P_X(\frac{y-b}{a}) = \begin{cases} 0 & \frac{y-b}{a} \leq -1 \\ \frac{1}{1-(2-1)} & -1 < \frac{y-b}{a} < 1 \\ 0 & x \geq 1 \end{cases} = \begin{cases} 0 & \frac{y-b}{a} \leq -1 \\ \frac{1}{1-(2-1)} & -1 < \frac{y-b}{a} < 1 \\ 0 & x \geq 1 \end{cases}$$

Section 3:
General Transformations

Section 4:
Monte Carlo Approximation

Section 5:
Entropy

Section 6:
Mutual Information