Machine Learning Foundations

Chapter 6: Probability

Harish Guruprasad Ramaswamy
IIT Madras

Outline for Chapter 6 : Probability

6.1 : Discrete Random Variables

6.2 : Continuous Random Variables

6.3 : Maximum Likelihood and other advanced topics

Outline for Chapter 6 : Probability

6.1 : Discrete Random Variables

1. Probability space
2. Conditioning
3. Random variables
4. Expectation and Variance
5. **Multiple Random Variables**
6. Bernoulli, Binomial, Poisson and Geometric RVs

6.2 : Continuous Random Variables

6.3 : Maximum Likelihood and other advanced topics

Joint Distributions

$$X, Y : \Omega \to \mathbb{R}$$

$$f_{XY}(x, y) = P(X = x, Y = y)$$

$$= P(\{\omega \in \Omega : X(\omega) = x\} \cap \{\omega \in \Omega : Y(\omega) = y\})$$

$$F_{XY}(x, y) = P(X \leq x, Y \leq y)$$

Examples

DDTE,

$P(A) = \frac{|A|}{36}$

X : Result of first die

Y : Absolute value of difference

$X \backslash Y$	0	1	2	3	4	5
1	1/36	1/36	1/36	1/36	1/36	1/36
2	1/36	2/36	1/36	1/36	1/36	0
3	1/36	2/36	2/36	1/36	0	0
4	1/36	2/36	2/36	1/36	0	0
5	1/36	2/36	1/36	1/36	1/36	0
6	1/36	1/36	1/36	1/36	1/36	1/36

Marginal and Conditional Distributions

$$f_X(x) = P(X = x)$$

$$= \bigcup_y P(X = x, Y = y)$$

$$= \sum_y f_{XY}(x, y)$$

$$f_Y(y) = \sum_{x \in Range(X)} f_{XY}(x, y)$$

$$f_{X|Y}(x|y) = P(X = x \mid Y = y)$$

$$= \frac{f_{XY}(x, y)}{f_Y(y)}$$

Examples

DDTE

X : First face

Y : Abs diff

Joint f_{XY}

$X \backslash Y$	0	1	2	3	4	5	f_X
1	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
2	$1/36$	$2/36$	$1/36$	$1/36$	$1/36$	0	$1/6$
3	$1/36$	$2/36$	$2/36$	$1/36$	0	0	$1/6$
4	$1/36$	$2/36$	$2/36$	$1/36$	0	0	$1/6$
5	$1/36$	$2/36$	$1/36$	$1/36$	$1/36$	0	$1/6$
6	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
f_Y	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{8}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	

Examples

DDTE

X : First face

Y : Abs diff

Conditional: $f_{X|Y}(x|y)$

$$\sum_x f_{X|Y}(x|y) = 1$$

X \ Y	0	1	2	3	4	5
1	1/6	1/10	1/8	1/6	1/4	1/2
2	1/6	2/10	1/8	1/6	1/4	0
3	1/6	2/10	2/8	1/6	0	0
4	1/6	2/10	2/8	1/6	0	0
5	1/6	2/10	1/8	1/6	1/4	0
6	1/6	1/10	1/	1/6	1/4	1/2
	1	1	1	1	1	1

Examples

DDTE

X : First face

Y : Abs diff

Conditional $f_{Y|X}(y|x)$

X \ Y	0	1	2	3	4	5	
1							1
2							1
3							1
4							1
5							1
6							1

Independent Random Variables

$$X, Y : \Omega \to \mathbb{R}$$

$$f_{XY}(x,y) = f_X(x) \cdot f_Y(y)$$

$$\Updownarrow$$

$\forall\, a, b$ $\{X = a\}$ & $\{Y = b\}$ are independent

$$\Updownarrow$$

$\forall\, g, h$ $\quad E[g(X)h(Y)] = E[g(X)] \cdot E[h(Y)]$

Independent Random Variables

$$E[g(X)h(Y)] = \sum_{x,y} f_{XY}(x,y)\, g(x)\, h(y)$$

$$= \sum_{x,y} f_X(x)\, f_Y(y)\, g(x)\, h(y)$$

$$= \sum_{x} f_X(x)\, g(x) \cdot \sum_{y} f_Y(y)\, h(y)$$

$$= E[g(X)] \cdot E[h(Y)]$$

Examples

DDTE,

$P(A) = \frac{|A|}{36}$

X = First face

Y = Second face

$\Rightarrow X, Y$ are Independent

$X \backslash Y$	1	2	3	4	5	6	
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2	1/36	1/36	1/36	1/36	1/36	1/36	1/6
3	1/36	1/36	1/36	1/36	1/36	1/36	1/6
4	1/36	1/36	1/36	1/36	1/36	1/36	1/6
5	1/36	1/36	1/36	1/36	1/36	1/36	1/6
6	1/36	1/36	1/36	1/36	1/36	1/36	1/6
	1/6	1/6	1/6	1/6	1/6	1/6	

Examples

DDTE

X : First face

Y : Abs difference.

NOT Independent.

$X \backslash Y$	0	1	2	3	4	5	f_X
1	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
2	$1/36$	$2/36$	$1/36$	$1/36$	$1/36$	0	$1/6$
3	$1/36$	$2/36$	$2/36$	$1/36$	0	0	$1/6$
4	$1/36$	$2/36$	$2/36$	$1/36$	0	0	$1/6$
5	$1/36$	$2/36$	$1/36$	$1/36$	$1/36$	0	$1/6$
6	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
f_X	$\frac{6}{36}$	$\frac{10}{36}$	$\frac{8}{36}$	$\frac{6}{36}$	$\frac{4}{36}$	$\frac{2}{36}$	

Sum of Random Variables

$X, Y : \Omega \to \mathbb{R}$ are independent

$Z = X + Y$

$$f_Z(z) = P(Z = z)$$

$$= \sum_x P(X = x, Y = z - x)$$

$$= \sum_x f_{XY}(x, z - x)$$

$$= \sum_x f_X(x) \cdot f_Y(z - x)$$

Conditional Expectation

$$X : \Omega \to \mathbb{R}$$

$$A \subseteq \Omega$$

$E[X|A]$ defined.

$$Y : \Omega \to \mathbb{R}$$

$E[X|Y]$ can be viewed as a function of Y

X : Sum of Die throw

Y : First die

$$E[X|Y] = Y + 3.5$$

Covariance

$$X, Y : \Omega \to \mathbb{R}$$

$$\operatorname{cov}[X, Y] := E\Big[(X - EX)(Y - EY)\Big]$$

$$\operatorname{cov}[X, X] := \operatorname{Var}[X]$$

$\operatorname{cov}[X, Y] = 0 \iff X, Y$ are uncorrelated.

Covariance

$$\mathrm{cov}[X,Y] = E[(X-EX)(Y-EY)]$$

$$= E[XY - (EX)Y - X\cdot EY + EX\cdot EY]$$

$$= E[XY] - EX\cdot EY - EX\cdot EY + EX\cdot EY$$

$$= E[XY] - EX\cdot EY$$

If X, Y are Independent

$$E[XY] \overset{\Downarrow}{=} EX\cdot EY \Rightarrow \mathrm{cov}[X,Y] = 0$$

Example

DCTE : $\Omega = \{HH, HT, TH, TT\}$

$X = 1(\text{first toss is Heads})$

$Y = 1(\text{second toss is tails})$

$$f_X(x) = \begin{cases} 1/2 & \text{if } x = 0 \\ 1/2 & \text{if } x = 1 \end{cases} \qquad EX = 1/2$$

$$f_Y(y) = \begin{cases} 1/2 & \text{if } y = 0 \\ 1/2 & \text{if } y = 0 \end{cases} \qquad EY = 1/2$$

$$f_{XY}(x, y) = \begin{cases} 1/4 & \text{if } x \in \{0, 1\} \;\&\; y \in \{0, 1\} \\ \end{cases}$$

Example

$$X, Y \Rightarrow \operatorname{cov}[X, Y] = 0$$

But NOT Independent

$$f_{XY}(x, y) =$$

$X \backslash Y$	-1	0	1	
-1	1/8	1/8	1/8	3/8
0	1/8	0	1/8	2/8
1	1/8	1/8	1/8	3/8
	3/8	2/8	3/8	

$$EX = 0$$

$$EY = 0$$

$$E[XY] = 0$$

$$P(XY = 1) = 1/4$$

$$P(XY = -1) = 1/4$$