SOOKMYUNG WOMEN'S UNIVERSITY K-MOOC

■ 결합분포와 주변분포

⊙ 동전 세 번 던지기

- ○ X: 앞면의 수, Y: 앞면과 뒷면의 수의 차이

$$\Omega = \{\ HHH,\ HHT,\ HTH,\ THH,\ HTT,\ THT,\ TTH\ \ TTT\ \}$$

	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
	↓	↓	↓	↓	↓	↓	↓	↓
$X=$	3	2	2	2	1	1	1	0
$Y=$	3	1	1	1	1	1	1	3

○ 두 변수를 동시에 고려한 확률분포?

Y \ X	0	1	2	3
1	?	?	?	?
3	?	?	?	?

○ 결합분포(joint distribution)

- 두 개 이상의 확률변수들을 동시에 고려한 확률분포

- 두 이산확률변수 X와 Y에 대해

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

  - $f(x,y)$: 결합확률질량함수

- n개의 이산확률변수 $X_1,\ ...,\ X_n$에 대해

$$f(x_1, ..., x_n) = P(X_1 = x_1, ..., X_n = x_n)$$

SOOKMYUNG WOMEN'S UNIVERSITY K-MOOC

⦿ 동전 세 번 던지기

X	3	2	2	2	1	1	1	0
Y	3	1	1	1	1	1	1	3

Y \ X	0	1	2	3
1	0	$\frac{3}{8}$	$\frac{3}{8}$	0
3	$\frac{1}{8}$	0	0	$\frac{1}{8}$

- $0 \leq f(x,y) \leq 1, \quad \forall\, x,\, y$
- $\sum_x \sum_y f(x,y) = 1$

SOOKMYUNG WOMEN'S UNIVERSITY K-MOOC

- 두 연속확률변수 X와 Y에 대해, 결합확률밀도함수 $f(x,\ y)$는 $x,\ y$에서의 밀도를 나타내며 아래의 성질을 만족
  - $0 \le f(x,y)$, $\quad \forall\, x,\ y$
  - $\displaystyle\int_x \int_y f(x,y)\, dy\, dx = 1$ ⇦ 부피 = 1

⊙ $(X, Y) \sim U((0,u),(0,v))$

$$f(x,y) = \frac{1}{uv}, \quad 0 < x < u,\ 0 < y < v$$

SOOKMYUNG WOMEN'S UNIVERSITY K-MOOC

○ 주변분포(marginal distribution)

- 표본공간이 사건 $B_1, \ldots, B_n$로 분할될 때 사건 A의 확률은

$$P(A) = \sum_{i=1}^{n} P(A \cap B_i)$$

  - ○ 사건 A가 $X = x$, B_i가 $Y = y_i$라고 하면

$$P(A \cap B_i) = P(X = x,\ Y = y_i)$$

- $P(X=x)=P(A)=\sum_{i=1}^{n} P(A \cap B_i) = \sum_{i=1}^{n} P(X=x, Y=y_i)$

  $$\Rightarrow f_X(x)=\sum_y f(x,y), \quad f_Y(y)=\sum_x f(x,y)$$

  - $f_X(x)$: X의 주변확률질량함수
  - $f_Y(y)$: Y의 주변확률질량함수

- 연속확률변수: 주변확률밀도함수

  $$f_X(x)=\int f(x,y)\,dy, \quad f_Y(y)=\int f(x,y)\,dx$$

SOOKMYUNG WOMEN'S UNIVERSITY K-MOOC

⊙ 동전 세 번 던지기

Y \ X		x				합
		0	1	2	3	
y	1	0	3/8	3/8	0	3/4
	3	1/8	0	0	1/8	1/4
합		1/8	3/8	3/8	1/8	1

⊙ $(X, Y) \sim U((0,u),(0,v))$

$$f_X(x) = \int_0^v \frac{1}{uv} dy = \frac{1}{uv} y \Big|_0^v = \frac{1}{u}, \ 0 < x < u$$

SOOKMYUNG WOMEN'S UNIVERSITY K-MOOC

○ 독립 확률변수

- 사건 A와 B는 독립 $\Leftrightarrow P(A \cap B) = P(A)P(B)$

- 두 확률변수 X와 Y는 독립 $\Leftrightarrow$ 모든 x, y에 대해,

$$f(x, y) = f_X(x) f_Y(y)$$

- $X_1, \ldots, X_n$이 서로독립(상호독립) $\Leftrightarrow$ 모든 $x_1, \ldots, x_n$에 대해,

$$f(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$$

⦿ 동전 세 번 던지기

$$f(1,1) = 3/8 \neq f_X(1) f_Y(1) = (3/8)(3/4) = 9/32$$

⇨ X와 Y는 독립이 아님

⊙ $f(x,y) = \dfrac{xy}{36}$, $x = 1,2,3$, $y = 1,2,3$

$Y \backslash X$	1	2	3	f_Y
1	1/36	2/36	3/36	1/6
2	2/36	4/36	6/36	2/6
3	3/36	6/36	9/36	3/6
f_X	1/6	2/6	3/6	1

○ 모든 x, y에 대해 $f(x,y) = f_X(x) f_Y(y)$ 성립

- $f(x,y) = g(x)h(y)$이고 x와 y의 값이 별개인 경우 $\Rightarrow$ 독립

SOOKMYUNG WOMEN'S UNIVERSITY K-MOOC

■ 요약

○ 두 이산확률변수 X와 Y에 대해, 결합확률질량함수

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

- $0 \le f(x,y) \le 1\,, \quad \forall\, x,\, y$
- $\sum_x \sum_y f(x,y) = 1$

○ 주변확률질량함수

$$f_X(x) = \sum_y f(x,y)\,, \quad f_Y(y) = \sum_x f(x,y)$$

○ 두 확률변수 X와 Y는 독립 $\Leftrightarrow$ 모든 $x,\, y$ 에 대해,

$$f(x,\, y) = f_X(x)\, f_Y(y)$$