

5-2 두 확률변수의 분포



5-2-1. 두 확률변수의 분포

결합확률분포

1. 이산형 확률변수인 경우

확률변수 X 의 가능한 값이 $x_1, \dots, x_n$ 이고 확률변수 Y 의 가능한 값이 $y_1, \dots, y_m$ 이라면 (X, Y) 의 결합확률분포는 다음과 같이 정의한다.

$$P_r(X = x_i, Y = y_j) = P_{ij}, i = 1, 2, \dots, n, j = 1, 2, \dots, m$$

여기에서 함수 $P(X = x_i, Y = y_j) = p_{ij}$ 를 결합확률질량함수(joint probability mass function) 라고 정의하며, 모든 (i, j) 에 대해 다음이 성립

$$p_{ij} \geq 0, \sum_{i=1}^n \sum_{j=1}^m p_{ij} = 1$$

2. 연속형 확률변수인 경우

두 확률변수 X, Y 의 결합확률밀도함수(joint probability density function)를 $f(x, y)$ 라고 정의할 때, 모든 (x, y) 에 대하여 $f(x, y) \geq 0$ 이고, 다음이 성립

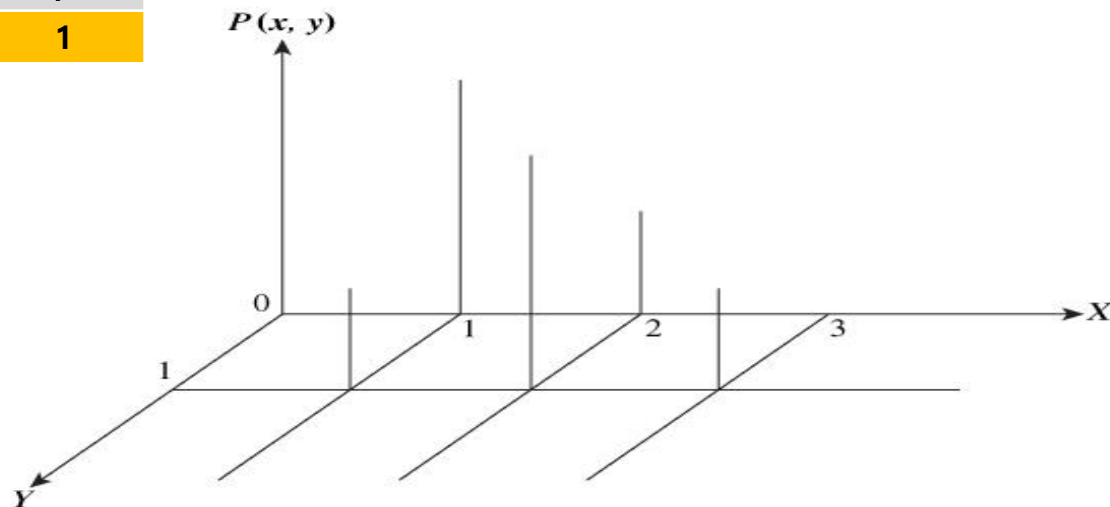
$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$$

예 동전 3개를 던지는 실험에서 확률변수 X 를 ' X =앞면의 수'라고 정의하고, 확률변수 Y 를 첫번째 동전이 앞면인 경우 $Y=1$, 첫번째 동전이 뒷면인 경우 $Y=0$ 이라고 정의한다. **X 와 Y 의 결합확률분포**를 표와 그래프는 나타내면 다음과 같다.

결과	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
X	3	2	2	2	1	1	1	0
Y	1	1	1	0	1	0	0	0
확률	1/8	1/8	1/8	1/8	1/8	1/8	1/8	1/8

X \ Y	0	1	합
0	1/8	0	1/8
1	2/8	1/8	3/8
2	1/8	2/8	3/8
3	0	1/8	1/8
합	1/2	1/2	1

그림 (X,Y)의 결합확률분포 그림



주변확률분포의 정의

1. 이산형 확률변수

X 의 가능한 값이 $(x_1, x_2, \dots, x_n)$, Y 의 가능한 값이 $(y_1, y_2, \dots, y_m)$,

X 와 Y 의 결합확률분포가 $P_r(X = x_i, Y = y_j) = P_{ij}, (i = 1, 2, \dots, n, j = 1, 2, \dots, m)$

이라고 할 때, X 와 Y 의 주변확률분포는 각각 다음과 같이 구한다.

$$X \text{의 주변확률분포} : P_r(X = x_i) = \sum_{j=1}^m p_{ij} = p_{i1} + p_{i2} + \dots + p_{im} \\ (i = 1, 2, \dots, n)$$

$$P_X(x) = \sum_j P_{X,Y}(x, y)$$

$$Y \text{의 주변확률분포} : P_r(Y = y_j) = \sum_{i=1}^n p_{ij} = p_{1j} + p_{2j} + \dots + p_{nj} \\ (j = 1, 2, \dots, m)$$

$$P_Y(y) = \sum_x P_{X,Y}(x, y)$$



2. 연속형 확률변수

두 확률변수 X, Y 의 결합확률밀도함수를 $f(x, y)$ 라고 할 때,
 X 와 Y 의 주변확률분포는 각각 다음과 같이 정의한다.

$$X \text{의 주변확률분포 : } f(x) = \int_{-\infty}^{\infty} f(x, y)dy$$

$$Y \text{의 주변확률분포 : } f(y) = \int_{-\infty}^{\infty} f(x, y)dx$$

연속형 확률분포에서 주변확률분포함수를

주변확률밀도함수(marginal probability density function)라고 부른다.



조건부확률분포의 정의

두 확률변수 X, Y 의 결합확률밀도함수를 $f(x, y)$ 라고 할 때,
어느 한 변수가 특정값을 취한다는 조건 하에 다른 변수가 값을 취할
확률을 나타내는 조건부확률분포는 각각 다음과 같이 정의한다.

$$f(x|y) = P(X=x|Y=y) \text{ (conditional PDF of X)}$$

$$f(y|x) = P(Y=y|X=x) \text{ (conditional PDF of Y)}$$

또는

$$f(x | y) = \frac{f(x, y)}{f(y)}, \quad f(y | x) = \frac{f(x, y)}{f(x)}$$

여기서 $f(x)$ 및 $f(y)$ 는 주변확률함수이다.



1. 이산형 확률변수

두 이산형 확률변수 X, Y 의 결합확률분포함수를

$P(X = x, Y = y)$ 라 하고, 각각의 주변분포함수를 $P(X = x)$, $P(Y = y)$ 라고 할 때, 모든 가능한 (x, y) 값에 대하여,

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

일 때 한하여 두 변수 X 와 Y 는 독립이다.



2. 연속형 확률변수

두 연속형 확률변수 X, Y 의 결합확률밀도함수를 $f(x, y)$ 라 하고, $f(x)$ 와 $f(y)$ 를 각각 X, Y 의 주변확률밀도함수라고 하면

$$f(x, y) = f(x) \cdot f(y)$$

일 때 한하여 두 변수 X 와 Y 는 독립이다.



예 실업률(X)과 경제성장률(Y)의 결합확률분포

X/Y	5%	10%	20%	X의 주변확률
3%	0.0	0.2	0.3	0.5
5%	0.1	0.1	0.2	0.4
10%	0.1	0.0	0.0	0.1
Y의 주변확률	0.2	0.3	0.5	1.0

$$p(X=3\% \mid Y=5\%) = ?$$

$$p(X=5\% \mid Y=5\%) = ?$$

$$p(X=10\% \mid Y=5\%) = ?$$

Question

실업률과 경제성장률은 서로 독립인가?