[3강]

통계학과 확률(2)

정보통계학과 이긍희 교수

학습목표

1. 결합확률분포를 이해한다.
2. 공분산과 상관계수를 이해한다.
3. 적률생성함수를 이해한다.
4. 조건부확률분포를 이해한다.

[3강]

통계학과 확률(2)

1 결합확률밀도(질량)함수

1.1 결합확률분포의 정의

- 여러 개의 확률변수가 동시에 관측
  → 여러 개 확률변수에 대한 분포는 결합확률밀도(질량)함수로 파악

1.1 결합확률분포의 정의

- 결합확률질량함수

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

$$f(x_1, \ldots, x_n) = P(X_1 = x_1, \ldots, X_n = x_n)$$

1.1 결합확률분포의 정의

- 결합확률밀도함수

$$P(a \le X \le b, c \le Y \le d) = \int_c^d \int_a^b f(x,y)dxdy$$

$$P(a_1 \le X_1 \le b_1, \ldots, a_n \le X_n \le b_n)$$

$$= \int_{a_n}^{b_n} \ldots \int_{a_1}^{b_1} f(x_1, \ldots, x_n)dx_1 \cdots dx_n$$

1.2 주변부확률밀도함수

	주변부확률질량함수	주변부확률밀도함수
$f_X(x)$	$\sum_{y=0}^{\infty} f(x,y)$	$\int_{-\infty}^{\infty} f(x,y)dy$
$f_Y(y)$	$\sum_{x=0}^{\infty} f(x,y)$	$\int_{-\infty}^{\infty} f(x,y)dx$

1.2 주변부확률밀도함수

예 2.16

동전 2개를 던져서 첫 번째 동전 앞면 여부를 X, 전체 동전 앞면의 수를 Y, 결합확률질량함수와 X, Y의 주변부확률질량함수는?

1.2 주변부확률밀도함수

예 2.17

$f(x,y) = 4xy, 0 \leq x, y \leq 1$일 때

$P\left(0 \leq X \leq \frac{1}{2}, \frac{1}{2} \leq Y \leq 1\right)$를 구하여라.

1.3 결합확률밀도(질량)함수의 성질

	결합확률질량함수	결합확률밀도함수
①	$0 \le f(x,y) \le 1$	$0 \le f(x,y)$
②	$\sum_{y=0}^{\infty}\sum_{x=0}^{\infty} f(x,y) = 1$	$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy = 1$
③	$P(a \le X \le b, c \le Y \le d) = \sum_{y=c}^{d}\sum_{x=a}^{b} f(x,y)$	$P(a \le X \le b, c \le Y \le d) = \int_{c}^{d}\int_{a}^{b} f(x,y)\,dxdy$

1.4 기댓값의 성질

- a, b가 상수이고 X, X_i, Y가 확률변수
  ① $E(aX + bY) = aE(X) + bE(Y)$
  ② $E(X_1 + \cdots + X_n) = E(X_1) + \cdots E(X_n)$

[3강]

통계학과 확률(2)

2 공분산

2.1 공분산

- **공분산(covariance)** : 두 확률변수가 선형적으로 같이 변하는 정도의 측도

$$Cov(X,Y) = E([X - E(X)][Y - E(Y)])$$

2.2 상관계수

- **상관계수(correlation)** : 표준화된 두 변수간 선형관계 척도

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

2.3 공분산과 상관계수의 성질

- 공분산과 상관계수의 성질

① $X = Y \Rightarrow Cov(X,Y) = Var(X)$,

$$Corr(X,Y) = 1$$

① $Cov(X,Y) = Cov(Y,X)$,

$Corr(X,Y) = Corr(Y,X)$

2.3 공분산과 상관계수의 성질

- 공분산과 상관계수의 성질

③ $Cov(aX+b, cY+d) = acCov(X,Y)$,

$$Corr(aX+b, cY+d) = \frac{ac}{|ac|}Corr(X,Y)$$

2.3 공분산과 상관계수의 성질

- 공분산과 상관계수의 성질

④ $Var(X \pm Y) = Var(X) + Var(Y)$
$\pm 2Cov(X,Y)$
$= Var(X) + Var(Y)$
$\pm 2Corr(X,Y)\sqrt{Var(X)Var(Y)}$

2.3 공분산과 상관계수의 성질

예 2.18

$Cov(X, Y)$와 $Corr(X, Y)$를 구하라.

y \ x	-1	1	합
-3	$\frac{1}{3}$	0	$\frac{1}{3}$
3	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
합	$\frac{2}{3}$	$\frac{1}{3}$	1

2.3 공분산과 상관계수의 성질

예 2.18

$Cov(X,Y)$와 $Corr(X,Y)$를 구하라.

2.4 확률변수 간 독립성

- $f_{X,Y}(x, y) = f_X(x) \times f_Y(y)$
- $f(x_1, \dots, x_n) = f_{X_1}(x_1) \times \cdots \times f_{X_n}(x_n)$

2.5 독립된 확률변수의 성질

- X, Y 독립 $\Rightarrow Cov(X,Y) = 0, Corr(X,Y) = 0$
  - $Var(X \pm Y) = Var(X) + Var(Y)$

2.5 독립된 확률변수의 성질

- $X_1, \dots, X_n$이 독립

$$Var(X_1 + \cdots + X_n) = Var(X_1) + \cdots + Var(X_n)$$

2.5 독립된 확률변수의 성질

예 2.20

$Cov(X,Y)$ 값을 구하고 X,Y가 독립인지 밝혀라.

y \ x	-1	0	1
0	0	$\frac{1}{3}$	0
1	$\frac{1}{3}$	0	$\frac{1}{3}$

2.5 독립된 확률변수의 성질

예 2.20

$Cov(X, Y)$ 값을 구하고 X, Y가 독립인지 밝혀라.

3 적률생성함수

3.1 적률생성함수

- **적률생성함수** : 모집단의 적률을 생성하는 함수

$$M(t) = E(e^{tX}) = \begin{cases} \displaystyle\sum_{x=0}^{\infty} e^{tx} f(x), \text{이산형} \\ \displaystyle\int_{-\infty}^{\infty} e^{tx} f(x) dx, \text{연속형} \end{cases}$$

3.1 적률생성함수

- 적률 : $E\left(X^k\right) = \frac{d^k}{dt^k} M(t) \Big]_{t=0}$

3.1 적률생성함수

예 2.21

X의 적률생성함수와 1차 적률인 기댓값은?

$$f(x) = \begin{cases} 2e^{-2x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

3.1 적률생성함수

예 2.21

X의 적률생성함수와 1차 적률인 기댓값은?

3.2 결합적률생성함수

- 두 확률변수 X, Y의 결합적률생성함수 :

$$M(t_1, t_2) = E(e^{t_1X+t_2Y})$$

- $$E(X^lY^m) = \frac{\partial^{l+m}}{\partial t_1^l \partial_2^m} M(t_1, t_2)\Big|_{t_1=t_2=0}$$

3.3 적률생성함수의 특성

- t가 0 부분에서 $M_X(t) = M_Y(t)$

$$\Leftrightarrow f_X(x) = f_Y(x)$$

- X, Y가 독립이면 $M_{X+Y}(t) = M_X(t)M_Y(t)$

[3강]

통계학과 확률(2)

4 조건부확률밀도함수

4.1 조건부확률밀도함수의 정의

- $X = x$ 하 Y의 조건부확률밀도함수$(f_X(x) > 0)$
  - 이산형 확률변수 :

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}, y = 0,1,2,\ldots$$

  - 연속형 확률변수 :

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}, -\infty < y < \infty$$

4.2 조건부확률밀도(질량)함수의 성질

	조건부확률질량함수	조건부확률밀도함수
①	$f_{Y\|X}(y\|x) \geq 0, y = 0,1,\ldots$	$f_{Y\|X}(y\|x) \geq 0, -\infty < y < \infty$
②	$\sum_{y=0}^{\infty} f_{Y\|X}(y\|x) = 1$	$\int_{-\infty}^{\infty} f_{Y\|X}(y\|x)\, dy = 1$
③	$P(a \leq y \leq b\|X = x) = \sum_{a\leq y\leq b} f_{Y\|X}(y\|x)$	$P(a \leq y \leq b\|X = x) = \int_{a}^{b} f_{Y\|X}(y\|x)\, dy$

4.2 조건부확률밀도(질량)함수의 성질

예 2.22

동전 2개를 던져서 첫 번째 동전 앞면 여부를 X, 전체 동전 앞면의 수를 Y. $X=1$일 때 Y의 조건부 확률질량함수는?

4.2 조건부확률밀도(질량)함수의 성질

예 2.23

두 확률변수 X, Y의 결합밀도함수가 다음과 같을 때 상수 C와 $f_{Y|X}(y|x)$를 각각 구하라.

$$f(x,y) = \begin{cases} C, & 0 < x < y < 2 \\ 0, & \text{그 밖의 구간} \end{cases}$$

4.2 조건부확률밀도(질량)함수의 성질

예 2.23

두 확률변수 X, Y의 결합밀도함수가 다음과 같을 때 상수 C와 $f_{Y|X}(y|x)$를 각각 구하라.

4.3 조건부 기댓값

$$E(Y|X=x) = \begin{cases} \sum_{y=0}^{\infty} y f_{Y|X}(y|x) & Y: \text{이산형} \\ \int_{-\infty}^{\infty} y f_{Y|X}(y|x)dy & Y: \text{연속형} \end{cases}$$

4.3 조건부 기댓값

예 2.24

$$f_{Y|X}(y|x) = \begin{cases} \frac{1}{1-x}, & x < y < 1 \\ 0, & \text{그밖의 구간} \end{cases}, \ E(Y|X = x)\text{는?}$$

4.4 조건부 기댓값의 성질

- a, b가 상수이고 X, Y가 확률변수

① $E(aY + bX|X = x) = aE(Y|X = x) + bx$

② $E[E(Y|X)] = E(Y)$

③ $Var(aY + bX|X = x) = a^2Var(Y|X = x)$

④ $Cov[Y - E(Y|X), E(Y|X)] = 0$

⑤ $Var(Y) = E\left(Var(Y|X)\right) + Var\left(E(Y|X)\right)$

[3강]

통계학과 확률(2)

5 정리하기

정리하기

1. 결합확률분포는 다음과 같이 정의된다.

이산형 : $f(x_1, \dots, x_n) = P(X_1 = x_1, \dots, X_n = x_n)$

연속형 : $P(a_1 \le X_1 \le b_1, \dots, a_n \le X_n \le b_n)$

$$= \int_{a_n}^{b_n} \dots \int_{a_1}^{b_1} f(x_1, \dots, x_n) dx_1 \cdots dx_n$$

정리하기

2. 두 확률변수가 같이 변하는 정도 측도로는 공분산(Cov)와 상관계수($Corr$)가 있다.

$$Cov(X,Y) = E[X - E(X)][Y - E(Y)]$$

$$Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}}$$

정리하기

3. 적률생섬함수 $M(t)$는 다음과 같이 정의되며 적률생성함수를 미분하여 적률을 구할 수 있다.

$$M(t) = E(e^{tX}) = \begin{cases} \displaystyle\sum_{x=0}^{\infty} e^{tx} f(x) \\ \displaystyle\int_{-\infty}^{\infty} e^{tx} f(x) dx \end{cases}$$

$$E(X^k) = \frac{d^k}{dt^k} M(t) \Big]_{t=0}$$

정리하기

4. $X = x$의 조건 하의 Y의 조건부확률밀도함수 $(f_X(x) > 0)$는 다음과 같다.

$$f_{Y|X}(y|x) = \frac{f(x,y)}{f_X(x)}$$

다음 시간 안내 ▼

4강. 모집단의 분포(1)

수고하셨습니다.