Chapter 5 다변량 확률분포함수

변수가 두 개 이상인 경우 이를 이변량, 다변량이라 한다.

○은행 고객이 기다리는 시간(X)과 서비스 받는 시간 (Y)

○바이러스 감염된 환자 수(X)와 24시간 후 치료된 환자 수 (Y)

○학생들의 키(X)와 몸무게(Y)

○전광판에 사용된 n개 전구의 수명($X_1, X_2, ..., X_n$)

이산형과 연속형 변수가 함께 있는 이변량, 다변량 분포도 가능하나 매우 복잡하고 유용성이 낮으므로 모두 연속형이거나 이산형인 경우에 한정하여 여기서는 다루기로 한다.

확률표본으로부터 계산되는 통계량(statistic, 예, 평균, 분산)은 ($X_1, X_2, ..., X_n$)의 함수이며 모집단의 모수를 추정(inference, 추정과 검정)하는데 사용된다. 그러므로 통계량의 확률밀도함수는 다변량 변수의 확률밀도함수이다. 우선 설명과 이해의 편의를 위하여 이변량 변수인 경우 확률밀도함수를 먼저 설명하기로 하자..

5.1 결합확률밀도함수(Joint PDF)

주사위를 동시에 두 개 던지는 실험에서 한 주사위의 눈금을 확률변수 X_1, 다른 주사위의 눈금을 X_2 라 정의하자. 이 확률실험에서 발생할 수 있는 경우의 수는 36가지이고 각각이 일어날 가능성(확률)이 동일하므로(equally likely) 확률변수 (X_1, X_2)의 결합밀도함수는 다음과 같이 정의할 수 있다.

$$p(x_1, x_2) = P(X_1 = x_1, X_2 = x_2) 1/36, \quad x_1 = 1,2,...,6; \quad x_2 = 1,2,...,6$$

정의 (joint probability distribution function) 결합확률밀도함수

$f(x_1, x_2) = P(X_1 = x_1, X_2 = x_2)$(이변량)
$f(x_1, x_2, ..., x_n) = P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n)$(다변량)

확률밀도함수이므로 다음이 성립한다. (다변량인 경우)

① $p(x_1, x_2, ..., x_n) \geq 0$ (이산형) $f(x_1, x_2, ..., x_n) \geq 0$ (연속형) for all $x_1, x_2, ..., x_n$

② $\sum_{x_1} ... \sum_{x_n} P(x_1, x_2, ..., x_n) = 1$ (이산형), $\int_{x_1} \int_{x_2} ... \int_{x_n} f(x_1, x_2, ..., x_n) dx_1 dx_2 ... dx_n = 1$ (연속형)

EXAMPLE 5.1

은행에 창구가 3개 있다. 2명의 고객이 서로 다른 시간에 왔고 서로 독립(모르는 사람)이라고 하자. 확률변수 X_1을 창구1에 온 고객 수, X_2을 창구 2에 온 고객 수라 할 때 (X_1, X_2) 결합밀도함수를 구하시오.

x2 \ x1	0	1	2
0	1/9	2/9	1/9
1	2/9	2/9	0
2	1/9	0	0

정의 (joint distribution function) 결합분포함수

$$F(x_1, x_2) = P(X_1 \le x_1, X_2 \le x_2) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f(x_1, x_2)dx_2 dx_1$$

$$F(x_1, x_2, ..., x_n) = P(X_1 \le x_1, X_2 \le x_2, ..., X_n \le x_n)$$

사건의 확률에서 $P(A \cap B) = P(AB)$과 동일하다.

EXAMPLE 5.2

Example 5.1에서 $F(-1,2)$, $F(1.5,2)$, $F(5,7)$을 구하시오.

0 / 8/9 / 1

정리(THEOREM) 이변량 결합분포함수의 성질

① $F(-\infty,-\infty) = F(-\infty, y_2) = F(y_1,-\infty) = 0$

② $F(\infty,\infty) = 1$

③비감소 함수(non-decreasing function)이다.

위의 정리의 증명은 obvious하다.

$P(a < X \le b, c < Y \le d) = F(b,d) - F(b,c) - F(a,d) + F(a,c)$ (그래프로 보자)

EXAMPLE 5.3

확률변수 (X_1, X_2) 의 결합확률밀도함수(joint PDF)는 $f(x_1, x_2) = 1, 0 \le x_1, x_2 \le 1$.

(1) $F(0.2, 0.4)$ 을 구하시오.

(2) $P(0.1 \le X_1 \le 0.3, 0 \le X_2 \le 5)$ 을 계산하시오.

0.08 / 0.1

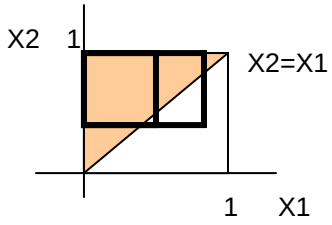
EXAMPLE 5.4

확률변수 (X_1, X_2) 의 결합확률밀도함수(joint PDF)는 $f(x_1, x_2) = k(1 - x_2), 0 \le x_1 \le x_2 \le 1$.

(1)상수 k 구하시오.

(2) $P(X_1 \le 3/4, X_2 \ge 1/2)$ 을 계산하시오.

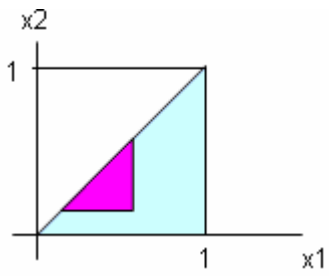
$k = 6$, 31/64



EXAMPLE 5.5

확률변수 X_1은 주유소 저장 탱크에 있는 GAS 양(%)이고 X_2는 하루에 팔린 GAS 양(%)이다. 결합밀도함수는 $f(x_1, x_2) = 3x_1, 0 \le x_2 \le x_1 \le 1$라 할 때 기름 탱크에 저장된 GAS 양이 1/2보다 작고 하루 팔린 양이 1/4 이하일 확률을 계산하시오.

5/128



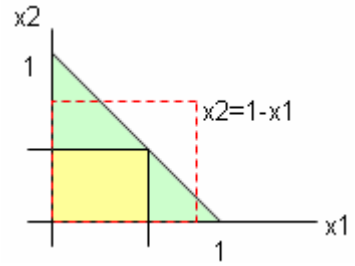
EXAMPLE 5.6

확률변수 (X_1, X_2) 의 결합밀도함수가 $f(x_1, x_2) = 2, 0 \le x_1 \le 1; 0 \le x_2 \le 1; 0 \le x_1 + x_2 \le 1$ 이다.

(1) $P(X_1 \le 1/2, X_2 \le 1/2)$ 을 계산하시오.

(2) $P(X_1 \le 3/4, X_2 \le 3/4)$ 을 계산하시오.

1/2, 10/16



HOMEWORK #12-1 DUE 5월 10일(화)

There are 9 executives in a company 4 are married, 3 are single, and 2 are divorced. Suppose that 3 executives are randomly selected. Let X_1 denote the number of married exec. and X_2 the number of never married. Find the joint pdf of (X_1, X_2)

HOMEWORK #12-2 DUE 5월 10일(화)

Let X_1 denote the total time at a bank between arrival and departure and X_2 the time a customer waits in line before reaching the service desk. $f(x_1, x_2) = e^{-x_1}, 0 \le x_2 \le x_1$

(1)Find $P(X_1 \le 2, X > 1)$.

(2)Find $P(X_1 \ge 2X_2)$.

(3)Find $P(X_1 - X_2 \ge 1)$ (the spent time at the service window)

HOMEWORK #12-3 DUE 5월 10일(화)

Let Y_1 and Y_2 denote the proportion of time during which employee I and II perform their assigned tasks, respectively. The joint pdf of (Y_1, Y_2) is $f(y_1, y_2) = y_1 + y_2, 0 \le y_1 \le 1; 0 \le y_2 \le 1$

(1)Find $P(Y_1 < 1/2, Y_2 > 1/4)$.

(2)Find $P(Y_1 + Y_2 \le 1)$.