2장. 다차원 확률변수의 확률분포

2.1 두 확률변수의 분포

예 2.1.1(p.53)(결합분포의 예)

두 개의 흰 공, 세 개의 검은 공, 네 개의 빨간 공이 들어있는 상자에서 3 개의 공을 함께 꺼낼 때, 추출되는 흰 공과 검은 공의 개수를 각각 X 와 Y 라고 하면

$$\mathrm{P}(\mathrm{X}=x, \mathrm{Y}=y)=\frac{\binom{2}{x}\binom{3}{y}\binom{4}{3-x-y}}{\binom{9}{3}},\ x=0,1,2,\ y=0,1,2,3,\ x+y\leq 3$$

		Y의 값				합계
		Y=0	Y=1	Y=2	Y=3	
X 의 값	X=0	4/84	18/84	12/84	1/84	35/84
	X=1	12/84	24/84	6/84	0	42/84
	X=2	4/84	3/84	0	0	7/84
합계		20/84	45/84	18/84	1/84	1

표 2.1.1 두 확률변수 X 와 Y 의 확률분포표

이산형 이변량 확률벡터와 확률밀도함수:(p.54)

(X, Y)가 가질 수 있는 순서쌍들의 집합: $\{(x_j, y_k) |\ j=1,2,\cdots, k=1,2,\cdots\}$

이차원[1] 확률변수 (X, Y)의 확률밀도함수: $f(x_j, y_k)=\mathrm{P}(\mathrm{X}=x_j, \mathrm{Y}=y_k)$

(a) $f(x,y)\geq 0 \quad \forall\, x,y: -\infty < x < +\infty,\ -\infty < y < +\infty$
$f(x,y)=0 \quad \forall\,(x,y)\neq(x_j, y_k)\ (j=1,2,\cdots,\ k=1,2,\cdots)$

(b) $\sum_x\sum_y f(x,y)=\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} f(x_j, y_k)=1$

(c) $\sum_{x: a\leq x\leq b}\ \sum_{y: c\leq y\leq d} f(x,y)=\mathrm{P}(a\leq \mathrm{X}\leq b,\ c\leq \mathrm{Y}\leq d)$

연속형 이차원 확률변수와 확률밀도함수:(p.54)

두 개의 확률변수 X, Y 가 모두 실수 구간의 값을 가질 수 있고 그에 관한 확률이 적분으로 주어질 때

(a) $f(x,y)\geq 0 \quad \forall\, x,y: -\infty < x < +\infty,\ -\infty < y < +\infty$

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

(c)[2] $\int_a^b\int_c^d f(x,y)\,dy\,dx=\mathrm{P}(a\leq \mathrm{X}\leq b,\ c\leq \mathrm{Y}\leq d) \quad (a<b,\ c<d)$

---

1) 이차원 확률변수를 이변량(bivariate) 확률벡터(random vector)라고도 한다.

2) 이중적분은 미분소인 $dy\,dx$ 가 나타나는 순서대로 적분해나가는 반복적분(iterated integral)을 뜻한다. 즉

$$\int_a^b\int_c^d f(x,y)\,dy\,dx=\int_a^b\left\{\int_c^d f(x,y)\,dy\right\}dx$$

부록 I : 미분과 적분

5. 이중합과 이중적분

정리 Ⅰ.5.1 토넬리(Tonelli) 정리[3](p.499)

(a:이중합)음이 아닌 함수 $f(x_1,x_2)$에 대하여

$$\sum_{x_1}\left\{\sum_{x_2} f(x_1,x_2)\right\} = \sum_{x_2}\left\{\sum_{x_1} f(x_1,x_2)\right\}$$

(b:이중적분)음이 아닌 함수 $f(x_1,x_2)$에 대하여

$$\int_{-\infty}^{+\infty}\left\{\int_{-\infty}^{+\infty} f(x_1,x_2)dx_2\right\}dx_1 = \int_{-\infty}^{+\infty}\left\{\int_{-\infty}^{+\infty} f(x_1,x_2)dx_1\right\}dx_2$$

일반적인 함수의 이중합과 이중 적분[4](p.499)

$$\sum_{x_1}\left\{\sum_{x_2} f(x_1,x_2)\right\} = \sum_{x_1}\left\{\sum_{x_2} f^+(x_1,x_2)\right\} - \sum_{x_1}\left\{\sum_{x_2} f^-(x_1,x_2)\right\}$$

$$\int_{-\infty}^{+\infty}\left\{\int_{-\infty}^{+\infty} f(x_1,x_2)dx_2\right\}dx_1 = \int_{-\infty}^{+\infty}\left\{\int_{-\infty}^{+\infty} f^+(x_1,x_2)dx_1\right\}dx_2 - \int_{-\infty}^{+\infty}\left\{\int_{-\infty}^{+\infty} f^-(x_1,x_2)dx_1\right\}dx_2$$

정리 Ⅰ.5.2 후비니(Fubini) 정리[5](p.499)

(a:이중합) 절대수렴하는 이중합의 경우에 다음이 성립한다.

$$\sum_{x_1}\left\{\sum_{x_2} f(x_1,x_2)\right\} = \sum_{x_2}\left\{\sum_{x_1} f(x_1,x_2)\right\}$$

(b:이중적분)적분가능 함수 $f(x_1,x_2)$에 대하여 다음이 성립한다.

$$\int_{-\infty}^{+\infty}\left\{\int_{-\infty}^{+\infty} f(x_1,x_2)dx_2\right\}dx_1 = \int_{-\infty}^{+\infty}\left\{\int_{-\infty}^{+\infty} f(x_1,x_2)dx_1\right\}dx_2$$

결합(結合 joint)분포와 주변(周邊 marginal)분포:(p.55)

(X, Y)의 분포는 결합(結合 joint)확률분포, X 의 분포와 Y의 분포는 각 확률변수의 주변(周邊 marginal)확률분포라고 한다.

결합확률밀도함수와 주변확률밀도함수:(p.56)

이차원 확률변수 (X, Y)의 결합확률밀도함수가 $f(x,y)$일 때, X 와 Y 의 주변확률밀도함수
$f_1(x)$와 $f_2(y)$는 각각 다음과 같이 주어진다.

$$f_1(x) = \begin{cases} \sum_y f(x,y) \\ \int_{-\infty}^{+\infty} f(x,y)dy \end{cases} \qquad f_2(y) = \begin{cases} \sum_x f(x,y) & \text{((X,Y)가 이산형일 때)} \\ \int_{-\infty}^{+\infty} f(x,y)dx & \text{((X,Y)가 연속형일 때)} \end{cases}$$

3)이 정리는 음이 아닌 함수를 전제로 하고 있기 때문에 이중합은 실수이거나 $+\infty$이다.
4)적분가능 함수의 정의에서와 같이 일반적인 함수의 이중합이나 이중적분은 절대수렴하는 것을 전제로 한다.
5)이 정리들은 다중적분의 경우에도 성립한다.

예 2.1.2 (반복 적분 계산 과정에서 연립부등식의 처리 예)(p.55)
다음 함수가 확률밀도함수가 되기 위한 상수 c 의 값을 구하고, 이 함수가 이차원 확률변수 (X, Y)의 확률밀도함수일 때 확률 $\mathrm{P}(\mathrm{X} \geq 2, \mathrm{Y} \geq 3)$을 구하여라.

$$f(x, y) = c e^{-x-y} \mathrm{I}_{(0 \leq x \leq y)}$$

[풀이] 전체 확률이 1 이어야 하므로

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,dy\,dx = \int_{0}^{+\infty}\int_{x}^{+\infty} c e^{-x-y}\,dy\,dx = 1$$

$$\therefore \int_{0}^{+\infty} c e^{-x}\left[-e^{-y}\right]_{x}^{+\infty} dx = c\int_{0}^{+\infty} e^{-2x}\,dx = c\left[-e^{-2x}/2\right]_{0}^{+\infty} = c/2 = 1$$

따라서 $c=2$ 이고 구하는 확률은

$$\mathrm{P}(\mathrm{X} \geq 2, \mathrm{Y} \geq 3) = \int_{2}^{+\infty}\int_{3}^{+\infty} 2e^{-x-y}\mathrm{I}_{(0 \leq x \leq y)}\,dy\,dx$$

그런데 $y \geq 3, y \geq x$을 간략히 $y \geq 3 \vee x$[6]로 나타내면

$$\int_{3}^{+\infty} 2e^{-x-y}\mathrm{I}_{(0 \leq x \leq y)}\,dy = \int_{3\vee x}^{+\infty} 2e^{-x-y}\,dy = 2e^{-x}\left[-e^{-y}\right]_{3\vee x}^{+\infty} = 2e^{-x-3\vee x}$$

$$\therefore \mathrm{P}(\mathrm{X} \geq 2, \mathrm{Y} \geq 3) = \int_{2}^{+\infty} 2e^{-x-3\vee x}\,dx = \int_{2}^{3} 2e^{-x-3}\,dx + \int_{3}^{+\infty} 2e^{-x-x}\,dx$$

$$\therefore \mathrm{P}(\mathrm{X} \geq 2, \mathrm{Y} \geq 3) = 2e^{-3}\left[-e^{-x}\right]_{2}^{3} + \left[-e^{-2x}\right]_{3}^{+\infty} = 2e^{-5} - e^{-6} \fallingdotseq 0.011$$

예 2.1.3 (이산형의 예)

예 2.1.4 (지표함수의 처리: 연립 부등식의 처리)(p.57)
예 2.1.2에서 (X, Y)의 결합확률밀도함수는

$$f(x, y) = 2e^{-x-y}\mathrm{I}_{(0 \leq x \leq y)}$$

이고, X 와 Y 의 주변확률밀도함수는 각각 다음과 같이 주어진다.

$$f_1(x) = \int_{-\infty}^{+\infty} 2e^{-x-y}\mathrm{I}_{(0 \leq x \leq y)}dy = \int_{x}^{+\infty} 2e^{-x-y}dy\,\mathrm{I}_{(0 \leq x)} = 2e^{-2x}\mathrm{I}_{(0 \leq x)}$$

$$f_2(y) = \int_{-\infty}^{+\infty} 2e^{-x-y}\mathrm{I}_{(0 \leq x \leq y)}dx = \int_{0}^{y} 2e^{-x-y}dx\,\mathrm{I}_{(0 \leq y)} = 2e^{-y}(1-\mathrm{e}^{-y})\mathrm{I}_{(0 \leq y)}$$

결합누적분포함수와 결합확률밀도함수:(p.58)
(a)(이산형인 경우)
X, Y 가 가질 수 있는 값들이 각각 $x_1 < x_2 < \cdots < x_m < \cdots$, $y_1 < y_2 < \cdots < y_n < \cdots$ 일 때,

$$\mathrm{F}(x_m, y_n) = \mathrm{P}(\mathrm{X} \leq x_m, \mathrm{Y} \leq y_n) = \sum_{j=1}^{m}\sum_{k=1}^{n} f(x_j, y_k)$$

$$f(x_j, y_k) = \mathrm{P}(\mathrm{X}=x_j, \mathrm{Y}=y_k) = \{\mathrm{F}(x_j, y_k) - \mathrm{F}(x_{j-1}, y_k)\} - \{\mathrm{F}(x_j, y_{k-1}) - \mathrm{F}(x_{j-1}, y_{k-1})\}$$

(b)(연속형인 경우)(p.58)

$$\mathrm{F}(x,y) = \mathrm{P}(\mathrm{X} \leq x\,\mathrm{Y} \leq y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(t,u)\,du\,dt$$

f 가 연속인 점 (x, y)에서 $f(x,y) = \dfrac{\partial^2}{\partial x \partial y}\mathrm{F}(x,y)$

6) $a \vee b$는 a와 b 중 큰 수를 나타내는 기호이고, $a \wedge b$는 a와 b 중 작은 수를 나타내는 기호이다.

주변누적분포함수:(p.58)

$$\mathrm{F}_1(x) = \mathrm{P}(\mathrm{X} \le x), \quad \mathrm{F}_2(y) = \mathrm{P}(\mathrm{Y} \le y)$$

를 각각 X 의 주변누적분포함수, Y의 주변누적분포함수라고 한다.

$$\mathrm{F}_1(x) = \lim_{y \to +\infty} \mathrm{F}(x, y), \quad \mathrm{F}_2(y) = \lim_{x \to +\infty} \mathrm{F}(x, y)$$

예 2.1.5 (지표함수와 연립부등식의 처리 예)(p.58)

예 2.1.2에서 (X, Y)의 결합누적분포함수는

$$\mathrm{F}(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} 2e^{-t-u} \mathrm{I}_{(0 \le t \le u)} du\, dt$$

그런데

$$\int_{-\infty}^{y} 2e^{-t-u} \mathrm{I}_{(0 \le t \le u)} du = 2e^{-t} \int_{t}^{y} e^{-u} du\, \mathrm{I}_{(y \ge t \ge 0)} = 2e^{-t}(e^{-t} - e^{-y}) \mathrm{I}_{(y \ge t \ge 0)}$$

$$\begin{aligned} \therefore\ \mathrm{F}(x,y) &= \int_{-\infty}^{x} 2e^{-t}(e^{-t} - e^{-y}) \mathrm{I}_{(0 \le t \le y)} dt \\ &= \int_{0}^{x \wedge y} 2(e^{-2t} - e^{-t-y}) dt\, \mathrm{I}_{(x \wedge y \ge 0)} \\ &= \left\{1 - e^{-2x \wedge y} + 2e^{-y}(e^{-x \wedge y} - 1)\right\} \mathrm{I}_{(x \ge 0,\, y \ge 0)} \end{aligned}$$

예 2.1.6 (지표함수의 처리 예)(p.59)

예 2.1.5에서 (X, Y)의 결합누적분포함수와 주변누적분포함수는

$$\mathrm{F}(x,y) = \left\{1 - e^{-2x \wedge y} + 2e^{-y}(e^{-x \wedge y} - 1)\right\} \mathrm{I}_{(x \ge 0,\, y \ge 0)}$$

$$\mathrm{F}_1(x) = \lim_{y \to +\infty} \left\{1 - e^{-2x \wedge y} + 2e^{-y}(e^{-x \wedge y} - 1)\right\} \mathrm{I}_{(x \ge 0)} = (1 - e^{-2x}) \mathrm{I}_{(x \ge 0)}$$

$$\mathrm{F}_2(y) = \lim_{x \to +\infty} \left\{1 - e^{-2x \wedge y} + 2e^{-y}(e^{-x \wedge y} - 1)\right\} \mathrm{I}_{(y \ge 0)} = (1 + e^{-2y} - 2e^{-y}) \mathrm{I}_{(y \ge 0)}$$

정리 2.1.1: 결합누적분포함수의 성질(p.59)

(a)(증가성)

(i) $\mathrm{F}(x,y) \le \mathrm{F}(x+h, y),\ \mathrm{F}(x,y) \le \mathrm{F}(x, y+k) \ \ \forall h > 0, k > 0$

(ii) $\mathrm{F}(x_2, y_2) - \mathrm{F}(x_1, y_2) - \mathrm{F}(x_2, y_1) + \mathrm{F}(x_1, y_1) \ge 0,\ \ \forall x_1 < x_2, \forall y_1 < y_2$

(b)(전체변동)

$$\lim_{x \to -\infty} \mathrm{F}(x, b) = 0, \quad \lim_{y \to -\infty} \mathrm{F}(a, y) = 0, \quad \lim_{\substack{x \to +\infty \\ y \to +\infty}} \mathrm{F}(x,y) = 1, \quad \forall a, b$$

(c)(오른쪽 연속성)

$$\lim_{\substack{h \to 0 \\ h > 0}} \mathrm{F}(x+h, y) = \mathrm{F}(x,y), \quad \lim_{\substack{k \to 0 \\ k > 0}} \mathrm{F}(x, y+k) = \mathrm{F}(x,y) \quad \forall x, y$$

2.2 결합확률분포의 특성치

결합확률분포에 대한 기댓값(p.60): 실수로 정의될 때

$$\mathrm{E}[g(\mathrm{X},\mathrm{Y})]=\begin{cases}\displaystyle\sum_x\sum_y g(x,y)pdf(x,y) & ((\mathrm{X},\mathrm{Y})\text{가 이산형일 때}) \\ \displaystyle\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g(x,y)pdf(x,y)\,dy\,dx & ((\mathrm{X},\mathrm{Y})\text{가 연속형일 때})\end{cases}$$

예 2.2.1 (연립부등식의 처리 예)(p.61)

이차원 확률변수 (X, Y)의 확률밀도함수가

$$f(x,y)=120\,xy(1-x-y)\mathrm{I}_{(x\,\ge\,0,\,y\,\ge\,0,\,x+y\,\le\,1)}$$

일 때, E(XY)를 구하여라.

[풀이]

$$\begin{aligned}\mathrm{E}(\mathrm{XY})&=\iint_{\substack{0\le x,0\le y\\ x+y\le 1}} xy\cdot 120xy(1-x-y)\,dy\,dx\\ &=\int_0^1\left\{120x^2\int_0^{1-x}y^2(1-x-y)\,dy\right\}dx\\ &=10\int_0^1 x^2(1-x)^4dx\\ &=10\int_0^1(1-t)^2t^4\,dt\\ &=2/21\end{aligned}$$

정리 2.2.1: 결합분포에 대한 기댓값의 성질[7)](p.61)

(a)(선형성) $\mathrm{E}[c_1g_1(\mathrm{X},\mathrm{Y})+c_2g_2(\mathrm{X},\mathrm{Y})]=c_1\mathrm{E}[g_1(\mathrm{X},\mathrm{Y})]+c_2\mathrm{E}[g_2(\mathrm{X},\mathrm{Y})]$ (c_1, c_2는 상수)

(b)(단조성) $g_1(\mathrm{X},\mathrm{Y})\le g_2(\mathrm{X},\mathrm{Y})$ 이면 $\mathrm{E}[g_1(\mathrm{X},\mathrm{Y})]\le\mathrm{E}[g_2(\mathrm{X},\mathrm{Y})]$

---

7) 이러한 선형성과 단조성은 각각의 기댓값이 실수로 정의되는 것을 전제로 하는 것이다.

공분산과 상관계수:(p.63)

(a)(공분산) 확률변수 X의 평균과 표준편차를 각각 μ_1과 σ_1, Y의 평균과 표준편차를 각각 μ_2와 σ_2라고 할 때, X와 Y의 공분산(共分散 covariance)

$$\mathrm{Cov}(X,Y) = E[(X-\mu_1)(Y-\mu_2)]$$

(b)(상관계수) 두 확률변수 X, Y의 표준편차 σ_1, σ_2 가 0이 아닐 때, X와 Y의 상관계수(相關係數 correlation coefficient)

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

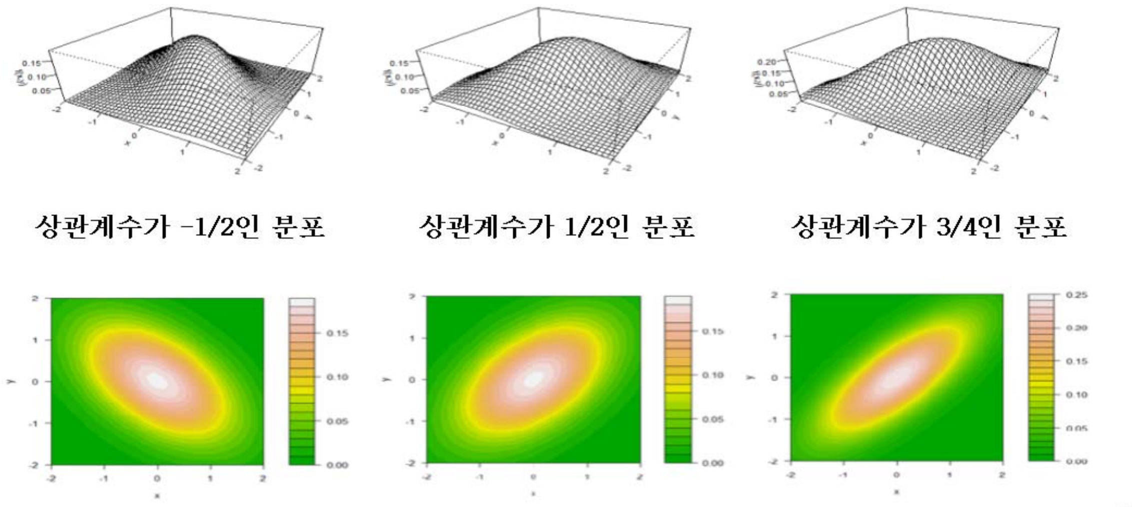

그림 2.2.1 여러 가지 분포의 모양과 상관계수의 값

정리 2.2.2: 공분산의 성질과 계산 공식[8](p.64)

(a) $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$, $\mathrm{Cov}(X,X) = \mathrm{Var}(X)$

(b) $\mathrm{Cov}(aX+b, cY+d) = ac\,\mathrm{Cov}(X,Y)$ $\quad (a,b,c,d$는 상수)

(c) $\mathrm{Cov}(X,Y) = E(XY) - [E(X)][E(Y)]$

[증명] (a)의 성질은 공분산의 정의로부터 명백하고, X, Y의 평균을 각각 μ_1, μ_2 라고 하면

$$E(aX+b) = a\mu_1 + b, \qquad E(cY+d) = c\mu_2 + d$$

(b)$V = aX+b$, $W = cY+d$ 라고 하면

$$V - E(V) = a(X-\mu_1),\ W - E(W) = c(Y-\mu_2)$$

$$\begin{aligned} \therefore\ \mathrm{Cov}(V,W) &= E[\{V-E(V)\}\{W-E(W)\}] \\ &= E[a(X-\mu_1)c(Y-\mu_2)] \\ &= ac\,\mathrm{Cov}(X,Y) \end{aligned}$$

(c) 공분산의 정의와 기댓값의 선형성으로부터

$$\begin{aligned} \mathrm{Cov}(X,Y) &= E[(X-\mu_1)(Y-\mu_2)] \\ &= E[(XY - \mu_1 Y - X\mu_2 + \mu_1\mu_2)] \\ &= E(XY) - 2\mu_1\mu_2 + \mu_1\mu_2 \\ &= E(XY) - [E(X)][E(Y)] \end{aligned}$$

---

8) 이러한 성질과 공식은 우변에 나타나는 기댓값과 공분산이 모두 실수라는 전제 하에 성립하는 것이다.

예 2.2.3(공분산 계산공식 적용 예)

예 2.2.1과 2.2.2에서 (X, Y)의 확률밀도함수가

$$f(x,y) = 120\,xy(1-x-y)\mathrm{I}_{(x\,\geq\,0,\,y\,\geq\,0,\,x+y\,\leq\,1)}$$

일 때,

$$\mathrm{E}(\mathrm{XY}) = 2/21, \quad \mathrm{E}(\mathrm{X}) = 1/3$$

같은 방법으로 $\mathrm{E}(\mathrm{Y}) = 1/3$ 임을 알 수 있고, 공분산의 계산 공식으로부터

$$\mathrm{Cov}(\mathrm{X},\mathrm{Y}) = \mathrm{E}(\mathrm{XY}) - [\mathrm{E}(\mathrm{X})][\mathrm{E}(\mathrm{Y})] = 2/21 - (1/3)(1/3) = -1/63$$

예 2.2.4(상관계수 계산의 예)

정리 2.2.3: 상관계수의 성질(p.65)

확률변수 X의 평균과 표준편차를 각각 μ_1과 $\sigma_1(\sigma_1 > 0)$, Y의 평균과 표준편차를 각각 μ_2와 $\sigma_2(\sigma_2 > 0)$, X와 Y의 상관계수를 ρ 라고 하면 다음이 성립한다.

(a) $$\mathrm{Var}\left(\frac{\mathrm{Y}-\mu_2}{\sigma_2} - \rho\frac{\mathrm{X}-\mu_1}{\sigma_1}\right) = 1-\rho^2$$

(b) $$-1 \leq \rho \leq 1$$

(c) $$\rho = 1 \Leftrightarrow \mathrm{P}\left(\frac{\mathrm{Y}-\mu_2}{\sigma_2} = \frac{\mathrm{X}-\mu_1}{\sigma_1}\right) = 1, \qquad \rho = -1 \Leftrightarrow \mathrm{P}\left(\frac{\mathrm{Y}-\mu_2}{\sigma_2} = -\frac{\mathrm{X}-\mu_1}{\sigma_1}\right) = 1$$

[증명] (a) $\mathrm{E}\left(\frac{\mathrm{Y}-\mu_2}{\sigma_2} - \rho\frac{\mathrm{X}-\mu_1}{\sigma_1}\right) = 0$ 이므로

$$\begin{aligned}\mathrm{Var}\left(\frac{\mathrm{Y}-\mu_2}{\sigma_2} - \rho\frac{\mathrm{X}-\mu_1}{\sigma_1}\right) &= \mathrm{E}\left[\left(\frac{\mathrm{Y}-\mu_2}{\sigma_2} - \rho\frac{\mathrm{X}-\mu_1}{\sigma_1}\right)^2\right] \\ &= \mathrm{E}\left[\left(\frac{\mathrm{Y}-\mu_2}{\sigma_2}\right)^2\right] - 2\rho\mathrm{E}\left[\left(\frac{\mathrm{Y}-\mu_2}{\sigma_2}\right)\left(\frac{\mathrm{X}-\mu_1}{\sigma_1}\right)\right] + \rho^2\mathrm{E}\left[\left(\frac{\mathrm{X}-\mu_1}{\sigma_1}\right)^2\right] \\ &= 1 - 2\rho\frac{\mathrm{Cov}(\mathrm{Y},\mathrm{X})}{\sigma_2\sigma_1} + \rho^2 \\ &= 1-\rho^2\end{aligned}$$

상관계수 부등식:

$$|\mathrm{Cov}(\mathrm{X},\mathrm{Y})| \leq \sqrt{\mathrm{Var}(\mathrm{X})}\sqrt{\mathrm{Var}(\mathrm{Y})}$$

상관계수는 두 변수 사이의 직선 관계를 나타내는 특성치:(p.66)

정리 2.2.3에서 알 수 있듯이 상관계수의 절댓값이 커질수록 (X, Y)의 분포는 직선

$$\frac{y-\mu_2}{\sigma_2} = \rho\frac{x-\mu_1}{\sigma_1}$$

주위에 가깝게 분포되어 나타날 것이다.

결합적률(p.66): (X, Y)의 $(r+s)$차의 (r,s)번째 결합적률(結合積率 joint moment)

$$\mathrm{E}(\mathrm{X}^r\mathrm{Y}^s)=\begin{cases}\sum_x\sum_y x^r y^s f(x,y) & ((\mathrm{X},\mathrm{Y})\text{가 이산형일 때})\\ \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x^r y^s f(x,y)\,dy\,dx & ((\mathrm{X},\mathrm{Y})\text{가 연속형일 때})\end{cases}$$

(지수함수의 멱급수 전개식)(p.67)

$$e^{t_1\mathrm{X}}e^{t_2\mathrm{Y}}=\sum_{r=0}^{\infty}\frac{(t_1\mathrm{X})^r}{r!}\sum_{s=0}^{\infty}\frac{(t_2\mathrm{Y})^s}{s!}=\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\frac{\mathrm{X}^r\mathrm{Y}^s}{r!s!}t_1^r t_2^s$$

(기댓값을 항별로 취하여 더할 수 있다면 다음을 예상)

$$\mathrm{E}(e^{t_1\mathrm{X}+t_2\mathrm{Y}})=\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\frac{\mathrm{E}(\mathrm{X}^r\mathrm{Y}^s)}{r!s!}t_1^r t_2^s \quad (???)$$

결합적률생성함수:(p.67)

0을 포함하는 열린구간들의 t_1,t_2 값에 대하여 $\mathrm{E}(e^{t_1\mathrm{X}+t_2\mathrm{Y}})$ 가 실수일 때, 함수

$$\mathrm{M}(t_1,t_2)=\mathrm{E}(e^{t_1\mathrm{X}+t_2\mathrm{Y}}),\ -h_1<t_1<h_1,-h_2<t_2<h_2\,(\exists\, h_1>0,h_2>0)$$

를 (X, Y)의 결합적률생성함수(joint moment generating function)라고 한다.

정리 2.2.4: 결합적률생성함수의 성질(p.67)

(a)(결합적률 생성)

$$\mathrm{M}(t_1,t_2)=\mathrm{E}(e^{t_1\mathrm{X}+t_2\mathrm{Y}})<+\infty,\ -h_1<t_1<h_1,\ -h_2<t_2<h_2\,(\exists\, h_1>0,h_2>0)$$

이면, (X, Y) 의 모든 결합적률이 존재하고

$$\mathrm{E}(\mathrm{X}^r\mathrm{Y}^s)=\left[\frac{\partial^{r+s}}{\partial t_1^r\partial t_2^s}\mathrm{M}(t_1,t_2)\right]_{\substack{t_1=0\\t_2=0}}\quad \mathrm{M}(t_1,t_2)=\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\frac{\mathrm{E}(\mathrm{X}^r\mathrm{Y}^s)}{r!s!}t_1^r t_2^s,\ -h_1<t_1<h_1,-h_2<t_2<h_2(\exists\, h_1>0,h_2>0)$$

(b)(분포 결정성)

$$\mathrm{M}_{\mathrm{X}_1,\mathrm{X}_2}(t_1,t_2)=\mathrm{M}_{\mathrm{Y}_1,\mathrm{Y}_2}(t_1,t_2)\ \ \forall\, t_i:-h_i<t_i<h_i\,(\exists\, h_i>0)(i=1,2)$$

이면, $(\mathrm{X}_1,\mathrm{X}_2)$와 $(\mathrm{Y}_1,\mathrm{Y}_2)$의 확률분포가 일치 즉 $(\mathrm{X}_1,\mathrm{X}_2)$와 $(\mathrm{Y}_1,\mathrm{Y}_2)$의 결합확률밀도함수, 결합누적분포함수가 일치한다.

결합누율생성함수(joint cumulant generating function)와 결합 누율:(p.68)

$$\mathrm{C}(t_1,t_2)=\log\mathrm{M}(t_1,t_2)=\log\mathrm{E}(e^{t_1\mathrm{X}+t_2\mathrm{Y}}),\ -h_i<t_i<h_i(\exists\, h_i>0)(i=1,2)$$

$$\mathrm{C}(t_1,t_2)=\sum_{\substack{r=0\\r+s\ge 1}}^{\infty}\sum_{s=0}^{\infty}\frac{\mathrm{C}^{(r,s)}(0,0)}{r!s!}t_1^r t_2^s,\quad \mathrm{C}^{(r,s)}(0,0)=\left[\frac{\partial^{r+s}}{\partial t_1^r\partial t_2^s}\mathrm{C}(t_1,t_2)\right]_{\substack{t_1=0\\t_2=0}}$$

에서 $t_1^r t_2^s/r!s!$의 계수인 $\mathrm{C}^{(r,s)}(0,0)$를 (X, Y)의 $(r+s)$차의 (r,s)번째 결합누율 이라고 하고 기호로는 $c_{r,s}(\mathrm{X},\mathrm{Y})$ 또는 $c_{r,s}$로 나타낸다.

결합누율과 결합 적률:(p.68)

$$\log(1+\sum_{\substack{j=0\\ j+k\ge 1}}^{\infty}\sum_{k=0}^{\infty}\frac{m_{j,k}}{j!k!}t_1^j t_2^k)=\sum_{\substack{r=0\\ r+s\ge 1}}^{\infty}\sum_{s=0}^{\infty}\frac{c_{r,s}}{r!s!}t_1^r t_2^s,\quad -h_i<t_i<h_i(\exists\, h_i>0)(i=1,2)$$

$$\log(1+\mathrm{A})=\mathrm{A}-\mathrm{A}^2/2+\mathrm{A}^3/3-\cdots,\ \ (-1<\mathrm{A}<1)$$

$$\mathrm{A}=(m_{1,0}t_1+m_{0,1}t_2)+(m_{2,0}t_1^2/2!+m_{1,1}t_1t_2+m_{0,2}t_2^2/2!)+\cdots$$

그 결과를 t_1, t_2의 오름차순으로 정리하면

$$c_{1,0}=m_{1,0}=\mathrm{E}(\mathrm{X})$$

$$c_{0,1}=m_{0,1}=\mathrm{E}(\mathrm{Y})$$

$$c_{2,0}=m_{2,0}-(m_{1,0})^2=\mathrm{E}(\mathrm{X}^2)-[\mathrm{E}(\mathrm{X})]^2=\mathrm{Var}(\mathrm{X})$$

$$c_{1,1}=m_{1,1}-m_{1,0}m_{0,1}=\mathrm{E}(\mathrm{XY})-[\mathrm{E}(\mathrm{X})][\mathrm{E}(\mathrm{Y})]=\mathrm{Cov}(\mathrm{X},\mathrm{Y})$$

$$c_{0,2}=m_{0,2}-(m_{0,1})^2=\mathrm{E}(\mathrm{Y}^2)-[\mathrm{E}(\mathrm{Y})]^2=\mathrm{Var}(\mathrm{Y})$$

예 2.2.5 (누율생성함수를 이용한 분산 공분산 계산의 예와 공식 확인 예)(p.69)

예 2.1.2에서 (X, Y)의 결합확률밀도함수는

$$f(x,y)=2e^{-x-y}\mathrm{I}_{(0\le x\le y)}$$

이므로, (X, Y)의 결합적률생성함수는

$$\mathrm{M}(t_1,t_2)=\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}e^{t_1x+t_2y}f(x,y)dydx=\int_0^{+\infty}\int_x^{+\infty}e^{t_1x+t_2y}2e^{-x-y}dy\,dx$$

$$\therefore \mathrm{M}(t_1,t_2)=\int_0^{+\infty}2e^{-(1-t_1)x}[-\frac{1}{1-t_2}e^{-(1-t_2)y}]_x^{+\infty}dx$$

$$=\frac{1}{1-t_2}\int_0^{+\infty}2e^{-(2-t_1-t_2)x}dx$$

$$=\frac{1}{1-t_2}\frac{2}{2-t_1-t_2},\ \ t_2<1, t_1+t_2<2$$

따라서 결합누율생성함수를 멱급수로 전개하면

$$\mathrm{C}(t_1,t_2)=-\log(1-t_2)-\log(1-(t_1+t_2)/2),\quad t_1+t_2<2, t_2<1$$

$$=\left\{t_2+t_2^2/2+\cdots\right\}+\left\{(t_1+t_2)/2+[(t_1+t_2)/2]^2/2+\cdots\right\}$$

$$=\left\{(1/2)t_1+(3/2)t_2\right\}+\left\{(1/4)t_1^2/2+(1/4)t_1t_2+(5/4)t_2^2/2\right\}+\cdots$$

그러므로

$$\mathrm{E}(\mathrm{X})=c_{1,0}=1/2,\ \ \mathrm{E}(\mathrm{Y})=c_{0,1}=3/2$$

$$\mathrm{Var}(\mathrm{X})=c_{2,0}=1/4,\ \mathrm{Cov}(\mathrm{X},\mathrm{Y})=c_{1,1}=1/4,\ \ \mathrm{Var}(\mathrm{Y})=c_{0,2}=5/4$$

이로부터

$$\mathrm{Corr}(\mathrm{X},\mathrm{Y})=\frac{\mathrm{Cov}(\mathrm{X},\mathrm{Y})}{\sqrt{\mathrm{Var}(\mathrm{X})}\sqrt{\mathrm{Var}(\mathrm{Y})}}=1/\sqrt{5}$$

2.3 조건부분포와 조건부기댓값

이산형인 경우의 조건부(條件附 conditional)확률밀도함수:(p.71)

$$f_{2|1}(y|x) = \frac{f_{1,2}(x,y)}{f_1(x)} = \mathrm{P}(\mathrm{Y}=y|\mathrm{X}=x) \quad (x : f_1(x) > 0)$$

(a) $f_{2|1}(y|x) \geq 0 \ \ \forall y : -\infty < y < +\infty$

$f_{2|1}(y|x) = 0 \ \ \forall y : y \neq y_k (k=1,2,\cdots)$

(b) $\sum_y f_{2|1}(y|x) = 1$

이차원 확률변수 (X, Y)가 연속형인 경우에는

$$\mathrm{P}(c \leq \mathrm{Y} \leq d, \mathrm{X}=x) = 0, \quad \mathrm{P}(\mathrm{X}=x) = 0$$

이므로, 1.2 절의 조건부확률의 정의로는 $\mathrm{X}=x$ 인 조건에서 Y에 관한 조건부확률을 생각할 수 없다. 그러나 양수 h에 대하여 조건부확률

$$\mathrm{P}(c \leq \mathrm{Y} \leq d | x \leq \mathrm{X} \leq x+h) = \frac{\mathrm{P}(c \leq \mathrm{Y} \leq d, x \leq \mathrm{X} \leq x+h)}{\mathrm{P}(x \leq \mathrm{X} \leq x+h)}$$

의 극한을 이용하여 생각할 수 있다. 즉 이산형의 경우와 같이, $f_1(x) > 0$ 일 때

$$f_{2|1}(y|x) = \frac{f_{1,2}(x,y)}{f_1(x)}$$

로 정의된 함수 $f_{2|1}(y|x)$에 대하여 다음 등식이 성립한다.

$$\lim_{\substack{h \to 0 \\ h > 0}} \mathrm{P}(c \leq \mathrm{Y} \leq d | x \leq \mathrm{X} \leq x+h) = \int_c^d f_{2|1}(y|x)\, dy \quad \text{(p.72)}$$

연속형인 경우의 조건부확률밀도함수와 조건부확률:(p.72)_

$$f_{2|1}(y|x) = \frac{f_{1,2}(x,y)}{f_1(x)} \quad (x : f_1(x) > 0)$$

(a) $f_{2|1}(y|x) \geq 0 \ \ \forall y : -\infty < y < +\infty$

(b) $\int_{-\infty}^{+\infty} f_{2|1}(y|x) dy = 1$

(c) $\mathrm{P}(c \leq \mathrm{Y} \leq d | \mathrm{X}=x) \equiv \int_c^d f_{2|1}(y|x)\, dy$ (조건부 확률의 정의)

예 2.3.2 (연속형 경우 조건부 확률 정의에 따른 계산의 예)(p.73)

예 2.1.2, 2.1.4에서 (X, Y)의 결합확률밀도함수와 X의 주변확률밀도함수는 각각

$$f_{1,2}(x,y) = 2e^{-x-y}\mathrm{I}_{(0 \leq x \leq y)}, \qquad f_1(x) = 2e^{-2x}\mathrm{I}_{(x \geq 0)}$$

따라서 $\mathrm{X}=x\,(x \geq 0)$인 조건에서 Y의 조건부확률밀도함수는

$$f_{2|1}(y|x) = e^{-(y-x)}\mathrm{I}_{(y \geq x)}$$

이고, 조건부확률은 그 정의에 따라 다음과 같이 주어진다.

$$\mathrm{P}(\mathrm{Y} \geq c | \mathrm{X}=x) = \int_c^{+\infty} e^{-(y-x)}\mathrm{I}_{(y \geq x)} dy = \int_{c \vee x}^{+\infty} e^{-(y-x)} dy = e^{-(c \vee x - x)}$$

정리 2.3.1: 조건부확률의 성질(p.73)

$$\mathrm{P}(a \le \mathrm{X} \le b,\ c \le \mathrm{Y} \le d) = \begin{cases} \sum_x \mathrm{P}(c \le \mathrm{Y} \le d | \mathrm{X} = x) f_1(x) & (\text{이산형인 경우}) \\ \int_{-\infty}^{+\infty} \mathrm{P}(c \le \mathrm{Y} \le d | \mathrm{X} = x) f_1(x) dx & (\text{연속형인 경우}) \end{cases}$$

[증명] (연속형의 경우)

$$\begin{aligned} \int_a^b \mathrm{P}(c \le \mathrm{Y} \le d | \mathrm{X} = x) f_1(x) dx &= \int_a^b \left\{ \int_c^d f_{2|1}(y|x) dy \right\} f_1(x) dx \\ &= \int_a^b \int_c^d f_{1,2}(x,y) dy\, dx \\ &= \mathrm{P}(a \le \mathrm{X} \le b,\ c \le \mathrm{Y} \le d) \end{aligned}$$

예 2.3.3 (정리 2.3.1을 확인하는 예)

예 2.1.2와 예 2.3.2에서, (X, Y)의 결합확률밀도함수가

$$f_{1,2}(x, y) = 2\, e^{-x-y} \mathrm{I}_{(0 \le x \le y)}$$

인 경우에 X의 주변확률밀도함수와 조건부확률은 각각 다음과 같다.

$$f_1(x) = 2e^{-2x} \mathrm{I}_{(x \ge 0)},\ \ \mathrm{P}(\mathrm{Y} \ge c | \mathrm{X} = x) = e^{-(c \vee x - x)}$$

따라서

$$\begin{aligned} \mathrm{P}(\mathrm{X} \ge 2, \mathrm{Y} \ge 3) &= \int_2^{+\infty} e^{-(3 \vee x - x)} 2e^{-2x} dx \\ &= \int_2^3 2e^{-3-x} dx + \int_3^{+\infty} 2e^{-2x} dx \\ &= 2e^{-5} - e^{-6} \end{aligned}$$

으로, 예2.1.2에서 구한 결과와 같다.

조건부평균(p.74):$\mathrm{X} = x$ 인 조건에서 Y의 조건부평균

$$\mu_{2|1}(x) = \mathrm{E}(\mathrm{Y} | \mathrm{X} = x) = \begin{cases} \sum_y y\, pdf_{2|1}(y|x) & (\text{이산형인 경우}) \\ \int_{-\infty}^{+\infty} y\, pdf_{2|1}(y|x)\, dy & (\text{연속형인 경우}) \end{cases}$$

예 2.3.4 (조건부 평균의 예)

예 2.2.1과 2.2.2에서, (X, Y)의 결합확률밀도함수와 X의 주변확률밀도함수가 각각

$$f_{1,2}(x, y) = 120\, xy(1 - x - y) \mathrm{I}_{(\mathrm{x} \ge 0,\, \mathrm{y} \ge 0,\, \mathrm{x} + \mathrm{y} \le 1)}, \qquad f_1(x) = 20x(1-x)^3 \mathrm{I}_{(0 \le x \le 1)}$$

이므로 $\mathrm{X} = x$ 인 조건에서 Y의 조건부확률밀도함수는 다음과 같이 주어진다.

$$f_{2|1}(y|x) = \frac{f_{1,2}(x,y)}{f_1(x)} = \frac{6y(1-x-y)}{(1-x)^3} \mathrm{I}_{(0 \le y \le 1-x)}$$

따라서 $\mathrm{X} = x$ 인 조건에서 Y의 조건부평균은

$$\mu_{2|1}(x) = \mathrm{E}(\mathrm{Y} | \mathrm{X} = x) = \int_0^{1-x} y \bullet 6y(1-x-y)/(1-x)^3 dy = (1-x)/2$$

조건부기댓값(p.75): $X=x$ 인 조건에서 $g(X,Y)$의 조건부기댓값

$$E(g(X,Y)|X=x)=\begin{cases}\sum_y g(x,y)\,pdf_{2|1}(y|x) & (\text{이산형인 경우})\\ \int_{-\infty}^{+\infty} g(x,y)\,pdf_{2|1}(y|x)\,dy & (\text{연속형인 경우})\end{cases}$$

조건부분산(p.75): $X=x$ 인 조건에서 Y의 조건부분산

$\mu_{2|1}(x)=E(Y|X=x)$ 라고 할 때,

$$Var(Y|X=x)=E[(Y-\mu_{2|1}(x))^2|X=x]=\begin{cases}\sum_y (y-\mu_{2|1}(x))^2\,pdf_{2|1}(y) & (\text{이산형인 경우})\\ \int_{-\infty}^{+\infty}(y-\mu_{2|1}(x))^2\,pdf_{2|1}(y)\,dy & (\text{연속형인 경우})\end{cases}$$

$Var(Y|X=x)$를 $\sigma^2_{2|1}(x)$로 나타내기도 한다.

정리 2.3.2: 조건부기댓값의 성질(p.76)

(a) (i) $E[c_1 g_1(Y)+c_2 g_2(Y)|X=x]=c_1 E[g_1(Y)|X=x]+c_2 E[g_2(Y)|X=x]$ (c_1, c_2는 상수)

(ii) $E[c(X)g(Y)|X=x]=c(x)E[g(Y)|X=x]$

(b) $g_1(Y)\le g_2(Y)$ 이면 $E[g_1(Y)|X=x]\le E[g_2(Y)|X=x]$

정리 2.3.3: 조건부분산의 계산공식(p.76)

$$Var(Y|X=x)=E(Y^2|X=x)-\{E(Y|X=x)\}^2$$

[증명] 조건부분산의 정의와 조건부기댓값의 성질로부터

$$\begin{aligned}Var(Y|X=x)&=E[(Y-\mu_{2|1}(x))^2|X=x]\\ &=E[Y^2-2\mu_{2|1}(x)Y+\mu^2_{2|1}(x)|X=x]\\ &=E(Y^2|X=x)-2\mu^2_{2|1}(x)+\mu^2_{2|1}(x)\\ &=E(Y^2|X=x)-\{E(Y|X=x)\}^2\end{aligned}$$

예 2.3.5 (조건부 분산의 예)(p.76)

예 2.3.4에서, $X=x$ 인 조건에서 Y의 조건부확률밀도함수와 조건부평균은 각각

$$f_{2|1}(y|x)=\frac{f_{1,2}(x,y)}{f_1(x)}=\frac{6y(1-x-y)}{(1-x)^3}I_{(0\le y\le 1-x)}$$

$$\mu_{2|1}(x)=E(Y|X=x)=(1-x)/2$$

따라서 $X=x$ 인 조건에서 Y의 조건부분산은

$$\begin{aligned}Var(Y|X=x)&=E(Y^2|X=x)-\{E(Y|X=x)\}^2\\ &=\int_0^{1-x} y^2\cdot 6y(1-x-y)/(1-x)^3dy-\{(1-x)/2\}^2\\ &=(1-x)^2/20\end{aligned}$$

확률변수로서의 조건부기댓값(p.77) $\mathrm{E}[g(\mathrm{Y})|\mathrm{X}]$:

$$h(x) = \mathrm{E}[g(\mathrm{Y})|\mathrm{X}=x]$$

라고 할 때, X의 분포에 따라 $h(x)$ 값들을 가지는 확률변수를

$$\mathrm{E}[g(\mathrm{Y})|\mathrm{X}] = h(\mathrm{X})$$

조건부평균(conditional mean of Y, given X): $\mathrm{E}(\mathrm{Y}|\mathrm{X}) = \mu_{2|1}(\mathrm{X})$

조건부분산(conditional variance of Y, given X): $\mathrm{Var}(\mathrm{Y}|\mathrm{X}) = \sigma_{2|1}^2(\mathrm{X})$

예 2.3.6 (확률변수로서 조건부 평균 분산의 예)

예 2.3.4와 2.3.5에서,

$$\mu_{2|1}(x) = \mathrm{E}(\mathrm{Y}|\mathrm{X}=x) = (1-x)/2, \quad \sigma_{2|1}^2(x) = \mathrm{Var}(\mathrm{Y}|\mathrm{X}=x) = (1-x)^2/20$$

따라서

$$\mathrm{E}(\mathrm{Y}|\mathrm{X}) = \mu_{2|1}(\mathrm{X}) = (1-\mathrm{X})/2, \quad \mathrm{Var}(\mathrm{Y}|\mathrm{X}) = \sigma_{2|1}^2(\mathrm{X}) = (1-\mathrm{X})^2/20$$

정리 2.3.4: 확률변수로서의 조건부기댓값의 성질[9](p.77)

(a) $\mathrm{E}[\mathrm{E}(\mathrm{Y}|\mathrm{X})] = \mathrm{E}(\mathrm{Y})$

(b) $\mathrm{Cov}(\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X}), v(\mathrm{X})) = 0, \ \forall v(\mathrm{X})$

[증명] (연속형인 경우)

(a)
$$\begin{aligned}\mathrm{E}[\mathrm{E}(\mathrm{Y}|\mathrm{X})] &= \int_{-\infty}^{+\infty} \mathrm{E}(\mathrm{Y}|\mathrm{X}=x) f_1(x)\,dx \\ &= \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} y f_{2|1}(y|x)\,dy \right\} f_1(x)\,dx \\ &= \int_{-\infty}^{+\infty} \left\{ \int_{-\infty}^{+\infty} y f_{2|1}(y|x) f_1(x)\,dy \right\} dx \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} y f_{1,2}(x,y)\,dy\,dx \\ &= \mathrm{E}(\mathrm{Y})\end{aligned}$$

(b) $\mathrm{Z} = \mathrm{Y} - \mathrm{E}(\mathrm{Y}|\mathrm{X})$라고 하면, 조건부기댓값의 성질과 (a)로부터

$$\mathrm{E}(\mathrm{Z}|\mathrm{X}) = \mathrm{E}(\mathrm{Y}|\mathrm{X}) - \mathrm{E}\{\mathrm{E}(\mathrm{Y}|\mathrm{X})|\mathrm{X}\} = \mathrm{E}(\mathrm{Y}|\mathrm{X}) - \mathrm{E}(\mathrm{Y}|\mathrm{X}) = 0, \ \mathrm{E}(\mathrm{Z}) = \mathrm{E}[\mathrm{E}(\mathrm{Z}|\mathrm{X})] = 0$$

따라서 공분산의 계산공식, 조건부기댓값의 성질과 (a)로부터

$$\begin{aligned}\mathrm{Cov}(\mathrm{Z}, v(\mathrm{X})) &= \mathrm{E}(\mathrm{Z}v(\mathrm{X})) - \mathrm{E}(\mathrm{Z})\mathrm{E}(v(\mathrm{X})) \\ &= \mathrm{E}\{\mathrm{E}(\mathrm{Z}v(\mathrm{X})|\mathrm{X})\} \\ &= \mathrm{E}\{v(\mathrm{X})\mathrm{E}(\mathrm{Z}|\mathrm{X})\} \\ &= 0\end{aligned}$$

예 2.3.7 (정리 2.3.4를 확인하는 예)(생략)

---

9) 이러한 성질은 (a) $\mathrm{E}(|\mathrm{Y}|) < +\infty$ (b)$\mathrm{E}(\mathrm{Y}^2) < +\infty, \mathrm{E}(v^2(X)) < +\infty$를 전제로 하는 것이다.

$\mathrm{E}(\mathrm{Y}|\mathrm{X})$: 회귀함수(回歸函數 regression function)(p.79)

정리 2.3.5: 최소제곱예측자(least squares predictor)(p.79)

$$\mathrm{E}[(\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X}))^2] \le \mathrm{E}[(\mathrm{Y}-u(\mathrm{X}))^2], \quad \forall\, u(\mathrm{X})$$

:평균제곱예측오차(mean squared prediction error) $\mathrm{E}[(\mathrm{Y}-u(\mathrm{X}))^2]$의 최솟값은

$$\mathrm{E}[(\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X}))^2] = \mathrm{E}[\mathrm{Var}(\mathrm{Y}|\mathrm{X})]$$

[Key] 예측 오차를 상관계수가 0인 두 확률변수의 합($\oplus$)으로 분해.

$$\mathrm{Y}-u(\mathrm{X}) = (\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X})) \oplus (\mathrm{E}(\mathrm{Y}|\mathrm{X})-u(\mathrm{X}))$$

여기에서 $v(\mathrm{X}) = \mathrm{E}(\mathrm{Y}|\mathrm{X}) - u(\mathrm{X})$라고 하면

$$\mathrm{E}[(\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X}))v(\mathrm{X})] = \mathrm{Cov}(\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X}), v(\mathrm{X})) = 0$$

$$\therefore\ \mathrm{E}[(\mathrm{Y}-u(\mathrm{X}))^2] = \mathrm{E}[(\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X}))^2] + \mathrm{E}[v^2(X)] \ge \mathrm{E}[(\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X}))^2]$$

한편, $\mathrm{W} = (\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X}))^2$ 라고 하면 조건부기댓값의 성질로부터

$$\mathrm{E}(\mathrm{W}|\mathrm{X}) = \mathrm{E}[(\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X}))^2|\mathrm{X}] = \mathrm{Var}(\mathrm{Y}|\mathrm{X})$$

$$\therefore\ \mathrm{E}[(\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X}))^2] = \mathrm{E}(\mathrm{W}) = \mathrm{E}[\mathrm{E}(\mathrm{W}|\mathrm{X})] = \mathrm{E}[\mathrm{Var}(\mathrm{Y}|\mathrm{X})]$$

정리 2.3.6: 분산의 분해

$$\mathrm{Var}(\mathrm{Y}) = \mathrm{E}[\mathrm{Var}(\mathrm{Y}|\mathrm{X})] + \mathrm{Var}[\mathrm{E}(\mathrm{Y}|\mathrm{X})]$$

[증명] Y의 평균을 μ라고 할 때,

$$\mathrm{Y}-\mu = (\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X})) \oplus (\mathrm{E}(\mathrm{Y}|\mathrm{X})-\mu)$$

$$\mathrm{E}[(\mathrm{Y}-\mu)^2] = \mathrm{E}[(\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X}))^2] + \mathrm{E}[(\mathrm{E}(\mathrm{Y}|\mathrm{X})-\mu)^2]$$

여기에서 $\mathrm{E}[\mathrm{E}(\mathrm{Y}|\mathrm{X})] = \mathrm{E}(\mathrm{Y}) = \mu$ 이므로

$$\mathrm{Var}(\mathrm{Y}) = \mathrm{E}[\mathrm{Var}(\mathrm{Y}|\mathrm{X})] + \mathrm{Var}[\mathrm{E}(\mathrm{Y}|\mathrm{X})]$$

예 2.3.8 (정리 2.3.6의 확인 예)(생략)

2.4 확률변수의 독립성

확률변수 X 와 Y 가 서로 독립(mutually independent):(p.81)
1)X 에 관한 어떠한 사건도 Y에 관한 사건과 서로 독립
2) a,b,c,d 의 값에 관계없이

$$\mathrm{P}(c \le \mathrm{Y} \le d | a \le \mathrm{X} \le b) = \mathrm{P}(c \le \mathrm{Y} \le d)$$

가 성립하는 경우로서, X 에 관한 어떠한 사건도 Y에 관한 사건과 서로 독립인 경우이다.
3) $\mathrm{P}(a \le \mathrm{X} \le b, c \le \mathrm{Y} \le d) = \mathrm{P}(a \le \mathrm{X} \le b)\mathrm{P}(c \le \mathrm{Y} \le d),\ \ \forall a,b,c,d(a<b, c<d)$

정리 2.4.1: 확률변수의 독립성[10]**(p.81)**
다음의 조건은 모두 확률변수 X 와 Y 가 서로 독립일 필요충분조건이다.
(a)(누적분포함수) $cdf_{1,2}(x,y) = cdf_1(x)cdf_2(y),\ \ \forall x,y : -\infty < x < +\infty, -\infty < y < +\infty$
(b)(확률밀도함수) $pdf_{1,2}(x,y) = pdf_1(x)pdf_2(y),\ \ \forall x,y : -\infty < x < +\infty, -\infty < y < +\infty$
(c)(적률생성함수)[11] $mgf_{1,2}(t_1,t_2) = mgf_1(t_1)mgf_2(t_2),\ \ \forall t_i : -h_i < t_i < h_i\,(\exists h_i > 0)(i=1,2)$
(d)(확률측도)[12] $\mathrm{P}(\mathrm{X}\in\mathrm{A}, \mathrm{Y}\in\mathrm{B}) = \mathrm{P}(\mathrm{X}\in\mathrm{A})\mathrm{P}(\mathrm{Y}\in\mathrm{B}),\ \ \forall \mathrm{A},\mathrm{B}$

예 2.4.1 (독립인 예)(확률밀도함수 이용 판단)
확률변수 X 와 Y 의 결합확률밀도함수:

$$f_{1,2}(x,y) = 2e^{-x-2y}\mathrm{I}_{(x \ge 0, y \ge 0)}$$

주변확률밀도함수:

$$f_1(x) = \int_{-\infty}^{+\infty} 2e^{-x-2y}\mathrm{I}_{(x \ge 0, y \ge 0)}dy = e^{-x}\mathrm{I}_{(x \ge 0)},\ \ f_2(y) = \int_{-\infty}^{+\infty} 2e^{-x-2y}\mathrm{I}_{(x \ge 0, y \ge 0)}dx = 2e^{-2y}\mathrm{I}_{(\mathrm{y} \ge 0)}$$

$$\therefore f_{1,2}(x,y) = 2e^{-x-2y}\mathrm{I}_{(x \ge 0)}\mathrm{I}_{(y \ge 0)} = f_1(x)f_2(y)\ \ \forall x,y$$

따라서 X 와 Y 는 서로 독립이다.

예 2.4.2 (독립이 아닌 예)(확률밀도함수 이용 판단)
예 2.1.2와 2.1.4에서 X 와 Y 의 결합확률밀도함수는

$$f_{1,2}(x,y) = 2e^{-x-y}\mathrm{I}_{(0 \le x \le y)}$$

이고, X 와 Y 의 주변확률밀도함수는 각각

$$f_1(x) = 2e^{-2x}\mathrm{I}_{(x \ge 0)},\qquad f_2(y) = 2e^{-y}(1-\mathrm{e}^{-y})\mathrm{I}_{(y \ge 0)}$$

이로부터

$$f_{1,2}(x,y) = f_1(x)f_2(y)\ \ \forall x,y$$

가 명백히 성립하지 않으므로, X 와 Y 는 서로 독립이 아니다.
 서로 독립이 아닌 확률변수는 **서로 종속(從屬 mutually dependent)**이라고 하며, 서로 종속인 관계를 밝히려면 정리 2.4.1의 항등식들이 성립하지 않는 하나의 예를 찾아야 한다.

---

10) cdf, pdf, mgf 는 각각 누적분포함수, 확률밀도함수, 적률생성함수를 뜻한다.
11) (c)의 조건은 적률생성함수가 존재할 때에 나머지 조건들과 동치이다.
12) $\mathrm{X}\in\mathrm{A}$ 는 $\{s : \mathrm{X}(s)\in\mathrm{A}\}$ 즉 역사상(inverse image)을 뜻한다.

정리 2.4.1의 (b)와 동등한 조건(p.83): (b')

(b') 다음 항등식이 성립하는 x만의 함수 $g_1(x)$와 y만의 함수 $g_2(y)$가 존재한다.

$$(pdf_{1,2}(x,y) = g_1(x)g_2(y) \ \forall \, x,y)$$

($\because$)

(b')이 성립하는 연속형의 경우에, 항등식의 양변을 모든 x와 y에 대하여 적분하면

$$\int_{-\infty}^{+\infty} g_1(x)dx \int_{-\infty}^{+\infty} g_2(y)dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} pdf_{1,2}(x,y)dydx = 1$$

이므로 다음 등식으로부터 (b)가 성립하는 것을 알 수 있다.

$$g_1(x)g_2(y) = \left\{ g_1(x) / \int_{-\infty}^{+\infty} g_1(x)dx \right\}\left\{ g_2(y) / \int_{-\infty}^{+\infty} g_2(y)dy \right\} = pdf_1(x)pdf_2(y)$$

(적용의 예)예 2.4.1의 경우에 주변확률밀도함수를 구하지 않고도

$$f_{1,2}(x,y) = 2e^{-x-2y}\mathrm{I}_{(x \geq 0, y \geq 0)} = 2e^{-x}\mathrm{I}_{(x \geq 0)}e^{-2y}\mathrm{I}_{(y \geq 0)} = g_1(x)g_2(y)$$

꼴이므로 X 와 Y 가 서로 독립이라고 할 수 있다.

정리 2.4.1의 (c)와 동등한 조건: (a)와 동등한 조건도 같은 형태로...

다음 항등식이 성립하는 t_1만의 함수 $m_1(t_1)$과 t_2만의 함수 $m_2(t_2)$가 존재한다.

$$mgf_{1,2}(t_1,t_2) = m_1(t_1)m_2(t_2)$$

정리 2.4.2: 독립인 확률변수들의 함수(p.83)

확률변수 X, Y가 서로 독립이면 각각의 함수인 $g_1(\mathrm{X})$와 $g_2(\mathrm{Y})$도 서로 독립이다.

[Key] 결합확률과 주변확률 사이의 관계 (d) 이용

함수 g의 역사상을 $g^{-1}(\mathrm{A}) = \{x : g(x) \in \mathrm{A}\}$로 나타내면

$$(g_1(\mathrm{X}) \in \mathrm{A}, g_2(\mathrm{Y}) \in \mathrm{B}) = (\mathrm{X} \in g_1^{-1}(\mathrm{A}), \mathrm{Y} \in g_2^{-1}(\mathrm{B}))$$

$$(g_1(\mathrm{X}) \in \mathrm{A}) = (\mathrm{X} \in g_1^{-1}(\mathrm{A})), \quad (g_2(Y) \in B) = (Y \in g_2^{-1}(B))$$

한편, X, Y가 서로 독립이면 정리 2.12의 (d)로부터

$$\mathrm{P}(\mathrm{X} \in g_1^{-1}(\mathrm{A}), \mathrm{Y} \in g_2^{-1}(\mathrm{B})) = \mathrm{P}(\mathrm{X} \in g_1^{-1}(\mathrm{A}))\mathrm{P}(\mathrm{X} \in g_2^{-1}(\mathrm{B}))$$

$$\therefore \mathrm{P}(g_1(\mathrm{X}) \in \mathrm{A}, g_2(\mathrm{Y}) \in \mathrm{B}) = \mathrm{P}(g_1(\mathrm{X}) \in \mathrm{A})\mathrm{P}(g_2(\mathrm{Y}) \in \mathrm{B})$$

따라서 $g_1(\mathrm{X})$와 $g_2(\mathrm{Y})$도 서로 독립이다.

정리 2.4.3: 독립인 확률변수들의 곱의 기댓값[13](p.84)

확률변수 X 와 Y 가 서로 독립이면

$$\mathrm{E}[g_1(\mathrm{X})g_2(\mathrm{Y})] = \mathrm{E}[g_1(\mathrm{X})]\mathrm{E}[g_2(\mathrm{Y})]$$

[증명] (연속형인 경우)

$$\begin{aligned}
\mathrm{E}[g_1(\mathrm{X})g_2(\mathrm{Y})] &= \int_{-\infty}^{+\infty}\left\{\int_{-\infty}^{+\infty} g_1(x)g_2(y)pdf_{1,2}(x,y)dy\right\}dx \\
&= \int_{-\infty}^{+\infty} g_1(x)pdf_1(x)\left\{\int_{-\infty}^{+\infty} g_2(y)pdf_2(y)dy\right\}dx \\
&= \int_{-\infty}^{+\infty} g_1(x)pdf_1(x)dx \int_{-\infty}^{+\infty} g_2(y)pdf_2(y)dy \\
&= E[g_1(\mathrm{X})]E[g_2(\mathrm{Y})]
\end{aligned}$$

---

13) 이 정리는 $g_1(\mathrm{X}), g_2(\mathrm{Y})$의 기댓값이 실수로 정의된다는 것을 전제로 한다.

정리 2.4.4: 독립성과 상관관계[14](p.84)
확률변수 X 와 Y 가 서로 독립이면

$$\mathrm{Cov}(X, Y) = 0$$

[증명] 공분산의 계산공식

$$\mathrm{Cov}(X,Y) = E(XY) - [E(X)][E(Y)]$$

과 정리 2.4.3으로부터 이 정리가 성립하는 것을 알 수 있다.

예 2.4.3 상관계수는 0 이지만 독립이 아닌 경우:(p.85)
확률변수 X 와 Y 의 결합확률밀도함수

$$f_{1,2}(x,y) = \frac{9}{40}(1+x^2y^2)\mathrm{I}_{(-1 \le x \le 1,\, -1 \le y \le 1)}$$

이면. 각각의 주변확률밀도함수는 다음과 같이 주어진다.

$$f_1(x) = \frac{3}{20}(3+x^2)\mathrm{I}_{(-1 \le x \le 1)}, \qquad f_2(y) = \frac{3}{20}(3+y^2)\mathrm{I}_{(-1 \le y \le 1)}$$

이 경우에 X 와 Y 는 명백히 서로 독립이 아니지만,

$$E(X) = 0,\ \ E(Y) = 0,\ \ E(XY) = 0\ \ \therefore\ \ \mathrm{Cov}(X,Y) = 0$$

즉 정리 2.4.4의 역이 항상 성립하는 것은 아니다.

정리 2.4.5: 확률변수의 합의 분산[15](p.85)
(a) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)$
(b) X 와 Y 가 서로 독립이면 $\mathrm{Var}(X \oplus Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$
[증명] X 와 Y 의 평균을 각각 μ_1, μ_2 라고 하면, 분산의 정의와 기댓값의 선형성으로부터

$$\begin{aligned}\mathrm{Var}(X+Y) &= E[(X+Y-\mu_1-\mu_2)^2] \\ &= E[(X-\mu_1)^2 + (Y-\mu_2)^2 + 2(X-\mu_1)(Y-\mu_2)] \\ &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y)\end{aligned}$$

---

14) 이 정리는 두 확률변수의 공분산이 실수로 정의된다는 것을 전제로 한다.
15) 이 정리는 각 확률변수의 분산이 실수로 정의된다는 것을 전제로 한다.

2.5 다차원 확률변수의 분포

이산형 다차원 확률변수의 결합확률밀도함수:(p.86)

(a) $f(x_1,\cdots,x_k) \ge 0 \quad \forall x_i : -\infty < x_i < +\infty \ (i=1,2,\cdots,k)$

$f(x_1,\cdots,x_k) = 0 \quad \forall x_i \not\in \{x_{i1}, x_{i2}, \cdots\} \ (i=1,2,\cdots,k)$

(b) $\sum_{x_1} \cdots \sum_{x_k} f(x_1,\cdots,x_k) = 1$

(c) $\sum_{x_1 : a_1 \le x_1 \le b_1} \cdots \sum_{x_k : a_k \le x_k \le b_k} f(x_1,\cdots x_k) = \mathrm{P}(a_1 \le \mathrm{X}_1 \le b_1, \cdots, a_k \le \mathrm{X_k} \le \mathrm{b_k})$

연속형 다차원 확률변수의 결합확률밀도함수:(p.86)

(a) $f(x_1,\cdots,x_k) \ge 0 \quad \forall x_i : -\infty < x_i < +\infty \ (i=1,2,\cdots,k)$

(b) $\int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x_1,\cdots,x_k)\, dx_k \cdots dx_1 = 1$

(c) $\int_{a_1}^{b_1} \cdots \int_{a_k}^{b_k} f(x_1,\cdots,x_k)\, dx_k \cdots dx_1 = \mathrm{P}(a_1 \le \mathrm{X}_1 \le b_1, \cdots, a_k \le \mathrm{X}_k \le b_k)$

예 2.5.1 (이산형 삼차원의 예)

예 2.5.2 (연속형 삼차원의 예)

$$f(x_1, x_2, x_3) = c e^{-x_1 - x_2 - x_3} \mathrm{I}_{(0 \le x_1 \le x_2 \le x_3)} \quad c = ?$$

주변확률밀도함수:(p.87)

다차원 확률변수 $(\mathrm{X}_1, \mathrm{X}_2, \cdots, \mathrm{X}_k)^{\mathrm{t}}$의 결합확률밀도함수가 $f(x_1,\cdots,x_k)$일 때, X_1과 $(\mathrm{X}_1, \mathrm{X}_2)^{\mathrm{t}}$의 주변확률밀도함수 $f_1(x)$와 $f_{1,2}(x,y)$는 각각 다음과 같이 주어진다.

(a)
$$f_1(x) = \begin{cases} \sum_{x_2} \cdots \sum_{x_k} f(x, x_2, \cdots, x_k) & (\text{이산형인 경우}) \\ \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x, x_2, \cdots, x_k) dx_k \cdots dx_2 & (\text{연속형인 경우}) \end{cases}$$

(b)
$$f_{1,2}(x,y) = \begin{cases} \sum_{x_3} \cdots \sum_{x_k} f(x, y, x_3, \cdots, x_k) & (\text{이산형인 경우}) \\ \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} f(x, y, x_3, \cdots, x_k) dx_k \cdots dx_3 & (\text{연속형인 경우}) \end{cases}$$

예 2.5.3 (이산형의 예)

예 2.5.1에서, $\mathrm{X}_1, \mathrm{X}_2, \mathrm{X}_3$ 중에서 $(\mathrm{X}_1, \mathrm{X}_2)^{\mathrm{t}}$ 의 주변확률밀도함수는

$$f_{1,2}(x,y) = \sum_{x_3} f(x,y,x_3) = \binom{2}{x}\binom{3}{y} \sum_{x_3} \binom{4}{x_3} \binom{5}{4-x-y-x_3} \Big/ \binom{14}{4}$$

$$\therefore \ f_{1,2}(x,y) = \binom{2}{x}\binom{3}{y}\binom{9}{4-x-y} \Big/ \binom{14}{4}, \ x = 0,1,2, \ y = 0,1,2,3, \ x+y \le 4$$

예 2.5.4 (연속형의 예)

예 2.5.2에서, $(X_1, X_2, X_3)^t$의 결합확률밀도함수가

$$f(x_1, x_2, x_3) = 6e^{-x_1 - x_2 - x_3} I_{(0 \le x_1 \le x_2 \le x_3)}$$

이므로 $(X_1, X_3)^t$ 의 주변확률밀도함수는

$$f_{1,3}(x,z) = \int_{-\infty}^{+\infty} f(x, x_2, z) dx_2 = 6e^{-x-z} \int_x^z e^{-x_2} dx_2 I_{(0 \le x \le z)}$$

$$\therefore\ f_{1,3}(x,z) = 6(e^{-2x-z} - e^{-x-2z}) I_{(0 \le x \le z)}$$

한편, X_1 의 주변밀도함수는

$$f_1(x) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, x_2, x_3) dx_3 dx_2 = 6\int_x^{+\infty} (e^{-2x-z} - e^{-x-2z}) dz I_{(0 \le x)}$$

$$\therefore\ f_1(x) = 6(e^{-2x} e^{-x} - e^{-x} e^{-2x}/2) I_{(0 \le x)} = 3e^{-3x} I_{(x \ge 0)}$$

다차원의 경우 확률밀도함수와 누적분포함수(연속형):(p.88)

$$F(x_1, \cdots, x_k) = P(X_1 \le x_1, \cdots, X_k \le x_k) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_k} f(t_1, \cdots t_k) dt_k \cdots dt_1$$

f 가 연속인 점 $(x_1, \cdots, x_k)$에서 $f(x_1, \cdots, x_k) = \dfrac{\partial^k}{\partial x_1 \cdots \partial x_k} F(x_1, \cdots, x_k)$

확률벡터의 함수의 기댓값(p.88): (절대수렴할 때)

$$E[g(X_1, \cdots, X_k)] = \begin{cases} \sum_{x_1} \cdots \sum_{x_k} g(x_1, \cdots, x_k) pdf(x_1, \cdots, x_k) & (\text{이산형인 경우}) \\ \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} g(x_1, \cdots, x_k) pdf(x_1, \cdots, x_k) dx_k \cdots dx_1 & (\text{연속형인 경우}) \end{cases}$$

예 2.5.5

확률벡터 $(X_1, X_2, X_3)^t$ 의 결합확률밀도함수가

$$f(x_1, x_2, x_3) = 120 x_1 (1 - x_1 - x_2 - x_3) I_{(X_1 \ge 0, X_2 \ge 0, X_3 \ge 0, X_1 + X_2 + X_3 \le 1)}$$

일 때, $E(X_1 X_2)$를 구하여라.

[풀이]

$$\begin{aligned} E(X_1 X_2) &= \iiint_{\substack{0 \le x, 0 \le y, 0 \le z \\ x+y+z \le 1}} xy \cdot 120x(1-x-y-z) dz\, dy\, dx \\ &= \int_0^1 \left\{ 120x^2 \int_0^{1-x} y \left\{ \int_0^{1-x-y} (1-x-y-z) dz \right\} dy \right\} dx \\ &= \int_0^1 \left\{ 120x^2 \int_0^{1-x} y(1-x-y)^2/2\, dy \right\} dx \\ &= \int_0^1 5x^2 (1-x)^4 dx \\ &= 1/21 \end{aligned}$$

정리 2.5.1: 기댓값의 성질[16](p.89)
(a)(선형성) 임의의 상수 c_1, c_2에 대하여

$$E[c_1 g_1(X_1,\cdots,X_k) + c_2 g_2(X_1,\cdots,X_k)] = c_1 E[g_1(X_1,\cdots,X_k)] + c_2 E[g_2(X_1,\cdots,X_k)]$$

(b)(단조성)

$$g_1(X_1,\cdots,X_k) \le g_2(X_1,\cdots,X_k) \text{ 이면 } \quad E[g_1(X_1,\cdots,X_k)] \le E[g_2(X_1,\cdots,X_k)]$$

확률벡터의 평균과 분산행렬:(p.90)
확률벡터[17] $X = (X_1,\cdots,X_k)^t$ 의 성분들인 $X_1,\cdots,X_k$의 평균, 분산 • 공분산인

$$\mu_i = E(X_i),\ \sigma_{i,j} = \mathrm{Cov}(X_i, X_j)\ \ (i,j = 1,2,\cdots,k)$$

를 대응하는 원소로 갖는 벡터와 행렬을 각각 X 의 평균벡터(mean vector), 분산 • 공분산행렬(variance-covariance matrix) 또는 간략히 X 의 평균, 분산행렬이라고 하며, 기호로는 E(X), Var(X)로 나타낸다. 즉,

$$E(X) = (\mu_1,\cdots,\mu_k)^t = (E(X_1),\cdots E(X_k))^t$$

$$\mathrm{Var}(X) = \begin{pmatrix} \sigma_{1,1} & \cdots & \sigma_{1,k} \\ \cdots & \cdots & \cdots \\ \sigma_{k,1} & \cdots & \sigma_{k,1} \end{pmatrix} = \left(\mathrm{Cov}(X_i, X_j)\right)_{\substack{1 \le i \le k \\ 1 \le j \le k}}$$

예 2.5.6(p.90)
예 2.2.1에서 확률벡터 $X = (X_1, X_2)^t$의 결합확률밀도함수가

$$f(x,y) = 120\,xy(1-x-y) I_{(x \ge 0,\, y \ge 0,\, x+y \le 1)}$$

일 때, 예 2.2.2, 2.2.3, 2.2.4 의 계산으로부터

$$E(X) = \begin{pmatrix} 1/3 \\ 1/3 \end{pmatrix}, \quad \mathrm{Var}(X) = \begin{pmatrix} 2/63 & -1/63 \\ -1/63 & 2/63 \end{pmatrix}$$

예 2.5.7
예 2.1.2에서 확률벡터 $X = (X_1, X_2)^t$의 결합확률밀도함수가

$$f(x,y) = 2\,e^{-x-y} I_{(0 \le x \le y)}$$

일 때, 예 2.2.5의 계산으로부터

$$E(X) = \begin{pmatrix} 1/2 \\ 3/2 \end{pmatrix}, \quad \mathrm{Var}(X) = \begin{pmatrix} 1/4 & 1/4 \\ 1/4 & 5/4 \end{pmatrix}$$

확률변수 $W_{i,j}(i,j = 1,\cdots,k)$를 성분으로 갖는 확률변수의 행렬 $(W_{i,j})$의 기댓값은 $E(W_{i,j})$를 대응하는 성분으로 갖는 행렬로 나타낸다. 즉,

$$E[(W_{i,j})] = (E(W_{i,j})), \qquad E\left[\begin{pmatrix} W_{1,1} & \cdots & W_{1,k} \\ \cdots & \cdots & \cdots \\ W_{k,1} & \cdots & W_{k,k} \end{pmatrix}\right] = \begin{pmatrix} E(W_{1,1}) & \cdots & E(W_{1,k}) \\ \cdots & \cdots & \cdots \\ E(W_{k,1}) & \cdots & E(W_{k,k}) \end{pmatrix}$$

이고, 이러한 확률변수 행렬의 기댓값도 다음과 같은 선형성을 갖는다.

---

16) 이러한 선형성과 단조성은 각각의 기댓값이 실수로 정의되는 것을 전제로 하는 것이다.
17) 벡터의 기호로는 $\underline{X}, \underline{a}, \underline{\mu}$와 같이 나타내어 실수 값의 X, a, μ와 구분하기도 하지만 이 책에서는 같은 기호를 사용하기로 한다. 따라서 문장이나 문단의 맥락에서 X, a, μ가 벡터 또는 실수인가를 판단하여야 한다.

정리 2.5.2: 확률변수 행렬의 기댓값의 성질[18]**(p.91)**

확률변수의 행렬 $\mathrm{V}=(\mathrm{V}_{i,j}),\mathrm{W}=(\mathrm{W}_{i,j})$에 대하여 다음이 성립한다.

(a) $\mathrm{E}(C\mathrm{W}D)=C\mathrm{E}(\mathrm{W})D$ (C,D는 모든 원소가 상수인 행렬)

(b) $\mathrm{E}(\mathrm{V}+\mathrm{W})=\mathrm{E}(\mathrm{V})+\mathrm{E}(\mathrm{W})$

분산행렬과 공분산행렬:(p.92)

$\mathrm{X}=(\mathrm{X}_1,\cdots,\mathrm{X}_k)^t$, $\mathrm{Y}=(\mathrm{Y}_1,\cdots,\mathrm{Y}_l)^t$ 의 평균을 각각 $\mu=(\mu_1,\cdots,\mu_k)^t$, $\eta=(\eta_1,\cdots,\eta_l)^t$ 라고 하면 X의 분산행렬과 X와 Y의 공분산행렬은 각각

$$\mathrm{Var}(\mathrm{X})=\mathrm{E}[(\mathrm{X}-\mu)(\mathrm{X}-\mu)^t],\quad \mathrm{Cov}(\mathrm{X},\mathrm{Y})=\mathrm{E}[(\mathrm{X}-\mu)(\mathrm{Y}-\eta)^t]$$

정리 2.5.3: 평균벡터와 공분산행렬의 성질[19]**(p.92)**

(a) $\mathrm{E}(A\mathrm{X}+b)=A\mathrm{E}(\mathrm{X})+b$ (A,b는 상수의 행렬과 벡터)

(b) $\mathrm{Var}(A\mathrm{X}+b)=A\mathrm{Var}(\mathrm{X})A^t$ (A,b는 상수의 행렬과 벡터)

(c) $\mathrm{Cov}(A\mathrm{X}+b,C\mathrm{Y}+d)=A\mathrm{Cov}(\mathrm{X},\mathrm{Y})C^t$ (A,C,b,d는 상수의 행렬과 벡터)

(d) $\mathrm{Cov}(\mathrm{X}+\mathrm{Y},\mathrm{Z})=\mathrm{Cov}(\mathrm{X},\mathrm{Z})+\mathrm{Cov}(\mathrm{Y},\mathrm{Z})$, $\mathrm{Cov}(\mathrm{X},\mathrm{Z}+\mathrm{W})=\mathrm{Cov}(\mathrm{X},\mathrm{Z})+\mathrm{Cov}(\mathrm{X},\mathrm{W})$

(e) $\mathrm{Cov}(\mathrm{Y},\mathrm{X})=(\mathrm{Cov}(\mathrm{X},\mathrm{Y}))^t$, $\mathrm{Var}(\mathrm{X})=\mathrm{Cov}(\mathrm{X},\mathrm{X})$

(f) $\mathrm{Var}(\mathrm{X}+\mathrm{Y})=\mathrm{Var}(\mathrm{X})+\mathrm{Var}(\mathrm{Y})+\mathrm{Cov}(\mathrm{X},\mathrm{Y})+\mathrm{Cov}(\mathrm{Y},\mathrm{X})$

[증명] 기댓값의 성질과 공분산행렬의 정의로부터 명백하다.

예 2.5.8(정리 2.5.3의 적용 예)

예 2.1.2에서 확률벡터 $\mathrm{X}=(\mathrm{X}_1,\mathrm{X}_2)^t$의 결합확률밀도함수가

$$f(x,y)=2e^{-x-y}\mathrm{I}_{(0\le x\le y)}$$

일 때, 예 2.5.7에서

$$\mathrm{E}(\mathrm{X})=\begin{pmatrix}1/2\\3/2\end{pmatrix},\quad \mathrm{Var}(\mathrm{X})=\begin{pmatrix}1/4&1/4\\1/4&5/4\end{pmatrix}$$

여기에서 $\mathrm{Y}=(\mathrm{Y}_1,\mathrm{Y}_2)^t=(2\mathrm{X}_1,\mathrm{X}_2-\mathrm{X}_1)^t$ 라고 하면

$$\mathrm{Y}=A\mathrm{X},\ A=\begin{pmatrix}2&0\\-1&1\end{pmatrix}$$

따라서 Y의 평균과 분산행렬은

$$\mathrm{E}(\mathrm{Y})=A\begin{pmatrix}1/2\\3/2\end{pmatrix}=\begin{pmatrix}1\\1\end{pmatrix},\quad \mathrm{Var}(\mathrm{Y})=A\begin{pmatrix}1/4&1/4\\1/4&5/4\end{pmatrix}A^t=\begin{pmatrix}1&0\\0&1\end{pmatrix}$$

정리 2.5.4 분산행렬의 성질:(nnd 행렬)(p.93)

확률벡터 $\mathrm{X}=(\mathrm{X}_1,\cdots,\mathrm{X}_k)^t$의 분산행렬 $\mathrm{Var}(\mathrm{X})$는 음이 아닌 정부호(nonnegative definite)의 대칭행렬이다. 즉

$$\mathrm{Var}(\mathrm{X})=(\mathrm{Var}(\mathrm{X}))^t,\qquad a^t\mathrm{Var}(\mathrm{X})a\ge 0,\ \ \forall a\in R^k$$

[증명] 분산행렬이 대칭행렬임은 명백하고, $a^t\mathrm{X}$은 실수 값을 가지는 확률변수이므로

$$a^t\mathrm{Var}(\mathrm{X})a=\mathrm{Var}(a^t\mathrm{X})\ge 0,\ \ \forall a\in R^k$$

---

18) 이러한 성질은 각각의 기댓값이 실수로 정의되는 것을 전제로 하는 것이다.

19) 이러한 성질은 각 확률변수의 분산이 실수로 정의되는 것을 전제로 하는 것이다.

--------------------------------(부록 II)--------------------------------

음이 아닌 정부호 행렬의 조건:(p.515)

실수가 원소인 $m \times m$대칭행렬 Σ에 대하여

$$a^t \Sigma a \geq 0, \ \forall a \in R^m$$

가 성립하면 Σ가 음이 아닌 정부호(nonnegative definite)의 행렬이라고 한다. 이러한 조건과 다음의 각 조건은 동등하다.

(a) Σ의 모든 고유값, 즉 $\det(\Sigma - \lambda I) = 0$인 λ는 음이 아닌 실수이다.

(b) $\Sigma = \Sigma^{1/2}\Sigma^{1/2}$ 인 실수가 원소인 대칭행렬 $\Sigma^{1/2}$가 존재한다

다차원 확률변수의 결합적률:(p.93)

$\mathrm{E}(|\mathrm{X}_1^{r_1} \cdots \mathrm{X}_k^{r_k}|) < +\infty$일 때

$$\mathrm{E}(\mathrm{X}_1^{r_1} \cdots \mathrm{X}_k^{r_k}) = \begin{cases} \sum_{x_1} \cdots \sum_{x_k} x_1^{r_1} \cdots x_k^{r_k} f(x_1,\cdots,x_k) & (\text{이산형인 경우}) \\ \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} x_1^{r_1} \cdots x_k^{r_k} f(x_1,\cdots,x_k) dx_k \cdots dx_1 & (\text{연속형인 경우}) \end{cases}$$

를 $\mathrm{X} = (\mathrm{X}_1,\cdots,\mathrm{X}_k)^t$의 $(r_1,\cdots r_k)$차 결합적률(結合積率 joint moment)이라고 하며. 기호로는 $m_{r_1,\cdots,r_k}(\mathrm{X}_1,\cdots,\mathrm{X}_k)$ 또는 $m_{r_1,\cdots,r_k}$로 나타낸다.

다차원 확률변수의 결합적률생성함수:(p.93)

0을 포함하는 열린구간들의 $t_1,\cdots,t_k$ 값에 대하여 $\mathrm{E}(e^{t_1\mathrm{X}_1 + \cdots + t_k\mathrm{X}_k})$ 가 실수일 때, 함수

$$\mathrm{M}(t_1,\cdots,t_k) = \mathrm{E}(e^{t_1\mathrm{X}_1 + \cdots + t_k\mathrm{X}_k}), \ -h_i < t_i < h_i, (\exists h_i > 0)(i = 1,\cdots,k)$$

를 $\mathrm{X} = (\mathrm{X}_1,\cdots,\mathrm{X}_k)^t$의 결합적률생성함수(joint moment generating function)라고 한다.

정리 2.5.5: 결합적률생성함수의 성질(p.94)

(a)(결합적률 생성) 다차원 확률변수 $\mathrm{X} = (\mathrm{X}_1,\cdots,\mathrm{X}_k)^t$의 결합적률생성함수가 존재하면, 즉

$$\mathrm{M}(t_1,\cdots,t_k) = \mathrm{E}(e^{t_1\mathrm{X}_1 + \cdots + t_k\mathrm{X}_k}) < +\infty, \ -h_i < t_i < h_i, (\exists h_i > 0)(i = 1,\cdots,k)$$

이면, $\mathrm{X} = (\mathrm{X}_1,\cdots,\mathrm{X}_k)^t$의 모든 결합적률이 존재하고

$$\mathrm{E}(\mathrm{X}_1^{r_1} \cdots \mathrm{X}_k^{r_k}) = \left[\frac{\partial^{r_1 + \cdots + r_k}}{\partial t_1^{r_1} \cdots \partial t_k^{r_k}} \mathrm{M}(t_1,\cdots,t_k)\right]_{t_1 = \cdots = t_k = 0}$$

$$\mathrm{M}(t_1,\cdots,t_k) = \sum_{r_1=0}^{\infty} \cdots \sum_{r_k=0}^{\infty} \frac{\mathrm{E}(\mathrm{X}_1^{r_1} \cdots \mathrm{X}_k^{r_k})}{r_1! \cdots r_k!} t_1^{r_1} \cdots t_k^{r_k}, \ -h_i < t_i < h_i (\exists h_i > 0) \ (i = 1,\cdots,k)$$

(b)(분포 결정성) 다차원 확률변수 $\mathrm{X} = (\mathrm{X}_1,\cdots,\mathrm{X}_k)^t$, $\mathrm{Y} = (\mathrm{Y}_1,\cdots,\mathrm{Y}_k)^t$ 의 결합적률생성함수 $\mathrm{M}_{\mathrm{X}_1,\cdots,\mathrm{X}_k}(t_1, \cdots,t_k)$, $\mathrm{M}_{\mathrm{Y}_1,\cdots,\mathrm{Y}_k}(t_1,\cdots,t_k)$가 존재하고 0을 포함하는 열린구간에서 일치하면, 즉

$$\mathrm{M}_{\mathrm{X}_1,\cdots,\mathrm{X}_k}(t_1, \cdots,t_k) = \mathrm{M}_{\mathrm{Y}_1,\cdots,\mathrm{Y}_k}(t_1,\cdots,t_k) \ \forall t_i : -h_i < t_i < h_i (\exists h_i > 0)(i = 1,\cdots,k)$$

이면, X와 Y의 확률분포가 일치 즉 $\mathrm{X} = (\mathrm{X}_1,\cdots,\mathrm{X}_k)^t$와 $\mathrm{Y} = (\mathrm{Y}_1,\cdots,\mathrm{Y}_k)^t$의 결합확률밀도함수, 결합누적분포함수가 일치한다.

결합누율생성함수와 평균, 분산:(p.95)

$$\mathrm{C}(t)=\log \mathrm{M}(t)=\log \mathrm{E}(e^{t_1 \mathrm{X}_1+\cdots+t_k \mathrm{X}_k}),\ -h_i<t_i<h_i(\exists\, h_i>0)(i=1,\cdots,k)$$

$\mathrm{C}(t)$의 일차 편도함수 벡터[20]와, 이차 편도함수 행렬[21]을 각각 $\dot{\mathrm{C}}(t)$, $\ddot{\mathrm{C}}(t)$라고 하면

$$\mathrm{C}(t)=\dot{\mathrm{C}}(0)^t t+\frac{1}{2}t^t\ddot{\mathrm{C}}(0)t+\cdots$$

$$\mathrm{E}(\mathrm{X})=\dot{\mathrm{C}}(0),\qquad \mathrm{Var}(\mathrm{X})=\ddot{\mathrm{C}}(0)$$

예 2.5.9(p.95)

예 2.5.2에서 확률벡터 $\mathrm{X}=(\mathrm{X}_1,\mathrm{X}_2,\mathrm{X}_3)^{\mathrm{t}}$의 결합확률밀도함수가

$$f(x_1,x_2,x_3)=6\,e^{-x_1-x_2-x_3}\mathrm{I}_{(0\le x_1\le x_2\le x_3)}$$

이므로, 결합적률생성함수를 구하면 다음과 같다.

$$\begin{aligned}\mathrm{M}_{\mathrm{X}}(t)&=\iiint_{0\le x_1\le x_2\le x_3} e^{t_1x_1+t_2x_2+t_3x_3}\bullet 6e^{-x_1-x_2-x_3}dx_3dx_2dx_1\\&=6\int_0^{+\infty}\left\{\int_{x_1}^{+\infty}[\int_{x_2}^{+\infty}e^{-(1-t_3)x_3}dx_3]e^{-(1-t_2)x_2}dx_2\right\}e^{-(1-t_1)x_1}dx_1\\&=6(3-t_1-t_2-t_3)^{-1}(2-t_2-t_3)^{-1}(1-t_3)^{-1},\ \ t_1+t_2+t_3<3,t_2+t_3<2,t_3<1\end{aligned}$$

따라서 결합누율생성함수는 $(t_1+t_2+t_3)/3<1,(t_2+t_3)/2<1,t_3<1$ 인 범위에서

$$\mathrm{C}(t)=-\log(1-\frac{t_1+t_2+t_3}{3})-\log(1-\frac{t_2+t_3}{2})-\log(1-t_3)$$

여기에서 로그함수의 멱급수 전개식

$$-\log(1-\mathrm{A})=\mathrm{A}+\mathrm{A}^2/2+\mathrm{A}^3/3+\cdots,\ \ (-1<\mathrm{A}<1)$$

을 이용하여 $\mathrm{C}(t)$를 멱급수 전개하고, 그 결과를 t_1,t_2,t_3의 오름차순으로 정리하면

$$\mathrm{C}(t)=\frac{1}{3}t_1+\frac{5}{6}t_2+\frac{11}{6}t_3+\frac{1}{2}(\frac{1}{9}t_1^2+\frac{13}{36}t_2^2+\frac{49}{36}t_3^2+\frac{1}{9}2t_1t_2+\frac{1}{9}2t_1t_3+\frac{13}{36}2t_2t_3)+\cdots$$

이 전개식의 일차항과 이차항의 계수로부터

$$\mathrm{E}(\mathrm{X})=\begin{pmatrix}1/3\\5/6\\11/6\end{pmatrix},\quad \mathrm{Var}(\mathrm{X})=\frac{1}{36}\begin{pmatrix}4&4&4\\4&13&13\\4&13&49\end{pmatrix}$$

---

20) 일차 편도함수 벡터 $\dot{\mathrm{C}}(t)=(\frac{\partial}{\partial t_1}\mathrm{C}(t),\cdots,\frac{\partial}{\partial t_k}\mathrm{C}(t))^t$ 를 기울기벡터(gradient vector)라고 한다.

21) 이차 편도함수 행렬 $\ddot{\mathrm{C}}(t)=\left(\frac{\partial^2}{\partial t_i\partial t_j}\mathrm{C}(t)\right)$ 을 헤시안행렬(Hessian matrix)이라고 한다.

다차원 확률변수에 대한 조건부확률밀도함수:(p.96)

다차원 확률변수 X, Y 에 대하여, X$=x$인 조건에서 Y의 조건부확률밀도함수는

$$f_{\mathrm{Y|X}}(y|x) = \frac{f_{\mathrm{X,Y}}(x,y)}{f_{\mathrm{X}}(x)} \quad (x : f_{\mathrm{X}}(x) > 0)$$

($\mathrm{X}_1 = x_1$인 조건에서 $(\mathrm{X}_2,\cdots,\mathrm{X}_k)^t$의 조건부확률밀도함수)

$$f_{2,\cdots,k|1}(x_2,\cdots,x_k|x_1) = \frac{f_{1,2,\cdots,k}(x_1,x_2,\cdots,x_k)}{f_1(x_1)} \quad (x_1 : f_1(x_1) > 0)$$

($(\mathrm{X}_2,\cdots,\mathrm{X}_k)^t = (x_1,\cdots,x_k)^t$인 조건에서 X_1의 조건부밀도함수)

$$f_{1|2,\cdots,k}(x_1|x_2,\cdots,x_k) = \frac{f_{1,2,\cdots,k}(x_1,x_2,\cdots,x_k)}{f_{2,\cdots,k}(x_2,\cdots,x_k)} \quad ((x_2,\cdots,x_k)^t : f_{2,\cdots,k}(x_2,\cdots,x_k) > 0)$$

다차원 확률변수의 조건부기댓값:(절대 수렴할 때)(p.97)

다차원 확률변수 $\mathrm{X} = (\mathrm{X}_1,\cdots,\mathrm{X}_k)^t, \mathrm{Y} = (\mathrm{Y}_1,\cdots,\mathrm{Y}_l)^t$에 대하여 $\mathrm{X} = x$ 인 조건에서 실수 값 함수 $g(\mathrm{X},\mathrm{Y})$의 조건부기댓값은

$$\mathrm{E}(g(\mathrm{X},\mathrm{Y})|\mathrm{X}=x) = \begin{cases} \sum_{y_1}\cdots\sum_{y_l} g(x,y_1,\cdots,y_l) f_{\mathrm{Y|X}}(y_1,\cdots,y_l|x) & (\text{이산형인 경우}) \\ \int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty} g(x,y_1,\cdots,y_l) f_{\mathrm{Y|X}}(y_1,\cdots,y_l|x)\, dy_l \cdots dy_1 & (\text{연속형인 경우}) \end{cases}$$

(<u>조건부평균과 조건부분산행렬</u>)

$$\mathrm{E}(\mathrm{Y|X}) = (\mathrm{E}(\mathrm{Y}_1|\mathrm{X}),\cdots,\mathrm{E}(\mathrm{Y}_l|\mathrm{X}))^t, \quad \mathrm{Var}(\mathrm{Y|X}) = (\mathrm{Cov}(\mathrm{Y}_i,\mathrm{Y}_j|\mathrm{X}))_{\substack{1 \le i \le l \\ 1 \le j \le l}}$$

예 2.5.10 (조건부확률밀도함수의 예)(p.97)

예 2.5.5에서 확률벡터 $(\mathrm{X}_1,\mathrm{X}_2,\mathrm{X}_3)^t$ 의 결합확률밀도함수가

$$f(x_1,x_2,x_3) = 120\,x_1(1-x_1-x_2-x_3)\mathrm{I}_{(\mathrm{x}_1 \ge 0, \mathrm{x}_2 \ge 0, \mathrm{x}_3 \ge 0, \mathrm{x}_1+\mathrm{x}_2+\mathrm{x}_3 \le 1)}$$

일 때, X_1과 $(\mathrm{X}_2,\mathrm{X}_3)^t$의 주변확률밀도함수는 각각

$$\begin{aligned} f_1(x_1) &= \iint_{\substack{x_2 \ge 0, x_3 \ge 0 \\ x_2+x_3 \le 1-x_1}} 120x_1(1-x_1-x_2-x_3)dx_3dx_2\mathrm{I}_{(0 \le x_1 \le 1)} \\ &= 20x_1(1-x_1)^3\mathrm{I}_{(0 \le \mathrm{x}_1 \le 1)} \end{aligned}$$

$$\begin{aligned} f_{2,3}(x_2,x_3) &= \int_0^{1-x_2-x_3} 120\,x_1(1-x_2-x_3-x_1)dx_1\mathrm{I}_{(x_2 \ge 0, x_3 \ge 0, x_2+x_3 \le 1)} \\ &= 20(1-x_2-x_3)^3\mathrm{I}_{(x_2 \ge 0, x_3 \ge 0, x_2+x_3 \le 1)} \end{aligned}$$

따라서, $\mathrm{X}_1 = x_1$인 조건에서 $(\mathrm{X}_2,\mathrm{X}_3)^t$의 조건부밀도함수는

$$f_{2,3|1}(x_2,x_3|x_1) = \frac{f_{1,2,3}(x_1,x_2,x_3)}{f_1(x_1)} = \frac{6(1-x_1-x_2-x_3)}{(1-x_1)^3}\mathrm{I}(x_2 \ge 0, x_3 \ge 0, x_2+x_3 \le 1-x_1)$$

이고, $(\mathrm{X}_2,\mathrm{X}_3)^t = (x_2,x_3)^t$인 조건에서 X_1의 조건부밀도함수는

$$f_{1|2,3}(x_1|x_2,x_3) = \frac{f_{1,2,3}(x_1,x_2,x_3)}{f_{2,3}(x_2,x_3)} = \frac{6x_1(1-x_2-x_3-x_1)}{(1-x_2-x_3)^3}\mathrm{I}_{(0 \le x_1 \le 1-x_2-x_3)}$$

예 2.5.11(조건부기댓값의 예)(p.98)

예 2.5.10에서, $X_1 = x_1$인 조건에서 $(X_2, X_3)^t$의 조건부밀도함수는

$$f_{2,3|1}(x_2, x_3|x_1) = \frac{f_{1,2,3}(x_1, x_2, x_3)}{f_1(x_1)} = \frac{6(1-x_1-x_2-x_3)}{(1-x_1)^3} I_{(x_2 \geq 0, x_3 \geq 0, x_2+x_3 \leq 1-x_1)}$$

$$\begin{aligned} E(X_2|X_1 = x_1) &= \int_0^{1-x_1}\int_0^{1-x_1-x_2} x_2 \cdot 6(1-x_1-x_2-x_3)dx_3dx_2/(1-x_1)^3 \\ &= (1-x_1)/4 \end{aligned}$$

$$\begin{aligned} E(X_2^2|X_1 = x_1) &= \int_0^{1-x_1}\int_0^{1-x_1-x_2} x_2^2 \cdot 6(1-x_1-x_2-x_3)dx_3dx_2/(1-x_1)^3 \\ &= (1-x_1)^2/10 \end{aligned}$$

$$\begin{aligned} E(X_2X_3|X_1 = x_1) &= \int_0^{1-x_1}\int_0^{1-x_1-x_2} x_2x_3 \cdot 6(1-x_1-x_2-x_3)dx_3dx_2/(1-x_1)^3 \\ &= (1-x_1)^2/20 \end{aligned}$$

$$Var(X_2|X_1 = x_1) = (1/10 - 1/16)(1-x_1)^2 = (3/80)(1-x_1)^2$$

$$Cov(X_2, X_3|X_1 = x_1) = (1/20 - 1/16)(1-x_1)^2 = (-1/80)(1-x_1)^2$$

$$\therefore\ E[\begin{pmatrix} X_2 \\ X_3 \end{pmatrix}|X_1] = \begin{pmatrix} 1/4 \\ 1/4 \end{pmatrix}(1-X_1),\quad Var[\begin{pmatrix} X_2 \\ X_3 \end{pmatrix}|X_1] = \frac{1}{80}\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}(1-X_1)^2$$

정리 2.5.6: 조건부기댓값의 성질(p.98)

(a) $E[E(Y|X)] = E(Y)$

(b) $Cov(Y - E(Y|X), v(X)) = 0,\ \forall v(X)$

정리 2.5.7: 분산행렬의 분해(p.99)

$$Var(Y) = E[Var(Y|X)] + Var[E(Y|X)]$$

[증명] 조건부기댓값의 성질로부터

$$Cov(Y - E(Y|X), E(Y|X)) = 0$$

$$\begin{aligned} Var(Y) &= Var[(Y - E(Y|X)) \oplus E(Y|X)] \\ &= Var(Y - E(Y|X)) + Var[E(Y|X)] \end{aligned}$$

한편, $W = Y - E(Y|X)$ 라고 하면,

$$E(WW^t|X) = E[(Y - E(Y|X))(Y - E(Y|X))^t|X] = Var(Y|X)$$

$$Var(W) = E(WW^t) = E[E(WW^t|X)] = E[Var(Y|X)]$$

$$\therefore\ Var(Y) = Var(W) + Var[E(Y|X)] = E[Var(Y|X)] + Var[E(Y|X)]$$

예 2.5.12(정리 2.5.7의 확인 예)

정리 2.5.8:다차원 확률변수의 최소제곱예측자(p.100)

다차원 확률변수 X의 벡터 값 함수 $u(\mathrm{X})=(u_1(\mathrm{X}),\cdots,u_k(\mathrm{X}))^t$로서

$$\mathrm{E}[\,\|\,\mathrm{Y}-u(\mathrm{X})\,\|^2]=\mathrm{E}[(\mathrm{Y}_1-u_1(\mathrm{X}))^2+\cdots+(\mathrm{Y}_k-u_k(\mathrm{X}))^2]$$

을 최소로 하는 벡터 값 함수는 $\mathrm{E}(\mathrm{Y}|\mathrm{X})=(\mathrm{E}(\mathrm{Y}_1|\mathrm{X}),\cdots,\mathrm{E}(\mathrm{Y}_k|\mathrm{X}))^t$이다. 즉

$$\mathrm{E}[\,\|\,\mathrm{Y}-\mathrm{E}(\mathrm{Y}|\mathrm{X})\,\|^2]\le \mathrm{E}[\,\|\,\mathrm{Y}-u(\mathrm{X})\,\|^2],\quad \forall\, u(\mathrm{X})$$

[증명] 조건부기댓값의 성질 (b)를 이용하여 정리 2.3.5와 같은 방법으로 증명할 수 있다.

다차원 확률변수의 독립성:(p.100)

$$\mathrm{P}(\mathrm{X}_1\in\mathrm{A}_1,\cdots,\mathrm{X}_n\in\mathrm{A}_n)=\mathrm{P}(\mathrm{X}_1\in\mathrm{A}_1)\cdots\mathrm{P}(\mathrm{X}_n\in\mathrm{A}_n),\quad \forall\,\mathrm{A}_i(i=1,\cdots,n)$$

가 성립할 때, 확률변수 $\mathrm{X}_1,\cdots,\mathrm{X}_n$이 서로 독립이라고 한다.

정리 2.5.9: 다차원 확률변수의 독립성[22]**(p.100)**

다음의 조건은 모두 다차원 확률변수 $\mathrm{X}_1,\cdots,\mathrm{X}_n$이 서로 독립일 필요충분조건이다.

(a)(확률밀도함수)

$$pdf_{1,\ldots,n}(x_1,\cdots,x_n)=pdf_1(x_1)\cdots pdf_n(x_n),\quad \forall\, x_i\,(i=1,\cdots,n)$$

(b)(적률생성함수)[23]

$$mgf_{1,\cdots,n}(t_1,\cdots,t_n)=mgf_1(t_1)\cdots mgf_n(t_n),\quad \forall\, t_i:\,\|\,t_i\,\|\,<h_i\,(\exists\, h_i>0)(i=1,\cdots,n)$$

예 2.5.13(p.101)

예 2.5.2에서, $(\mathrm{X}_1,\mathrm{X}_2,\mathrm{X}_3)^{\mathrm{t}}$의 결합확률밀도함수가

$$f_{1,2,3}(x_1,x_2,x_3)=6\,e^{-x_1-x_2-x_3}\mathrm{I}_{(0\le x_1\le x_2\le x_3)}$$

일 때 $\mathrm{X}_1,\mathrm{X}_2,\mathrm{X}_3$의 주변밀도함수는 각각 다음과 같이 주어진다.

$$f_1(x_1)=6e^{-x_1}\int_{x_1}^{+\infty}e^{-x_2}\int_{x_2}^{+\infty}e^{-x_3}dx_3dx_2=3\,e^{-3x_1}\mathrm{I}_{(x_1\ge 0)}$$

$$f_2(x_2)=6e^{-x_2}\int_0^{x_2}e^{-x_1}\int_{x_2}^{+\infty}e^{-x_3}dx_3dx_1=6\,e^{-2x_2}(1-e^{-x_2})\mathrm{I}_{(x_2\ge 0)}$$

$$f_3(x_3)=6e^{-x_3}\int_0^{x_3}e^{-x_1}\int_{x_1}^{x_3}e^{-x_2}dx_2dx_1=3\,e^{-x_3}(1-e^{-x_3})^2\mathrm{I}_{(x_3\ge 0)}$$

따라서 결합밀도함수가 주변밀도함수의 곱으로 나타내어지지 않으므로 $\mathrm{X}_1,\mathrm{X}_2,\mathrm{X}_3$는 서로 독립이 아니다.

한편, 두 확률변수가 서로 독립임을 밝힐 때와 마찬가지로 정리 2.5.9를 이용하여 여러 확률변수가 서로 독립임을 밝히는 과정에서 주변확률밀도함수나 주변적률생성함수를 반드시 구할 필요는 없다. 즉, 위의 (a), (b) 대신 다음의 (a'), (b')을 적용하여도 된다.

(a') 다음 항등식이 성립하는 함수 $g_1,\cdots,g_n$이 존재한다.

$$pdf_{1,\ldots,n}(x_1,\cdots,x_n)=g_1(x_1)\cdots g_n(x_n),\quad \forall\, x_i\,(i=1,\cdots,n)$$

(b') 다음 항등식이 성립하는 함수 $m_1,\cdots,m_n$이 존재한다.

$$mgf_{1,\cdots,n}(t_1,\cdots,t_n)=m_1(t_1)\cdots m_n(t_n),\quad \forall\, t_i:\,\|\,t_i\,\|\,<h_i\,(\exists\, h_i>0)(i=1,\cdots,n)$$

---

22) pdf, mgf 는 각각 확률밀도함수, 적률생성함수를 뜻한다.

23) (b)의 조건은 적률생성함수가 존재하는 것을 전제로 하는 것이다.

예 2.5.14(p.102)
예 2.5.9에서의 확률벡터 $X=(X_1,X_2,X_3)^t$에 대하여 $Y_1=X_1,\ Y_2=X_2-X_1,\ Y_3=X_3-X_2$
라고 할 때, Y_1,Y_2,Y_3 가 서로 독립임을 밝혀라.
[풀이] 예 2.5.9에서 $X=(X_1,X_2,X_3)^t$의 결합적률생성함수는

$$\begin{aligned} M_X(t_1,t_2,t_3) &= E(e^{t_1X_1+t_2X_2+t_3X_3}) \\ &= 6(3-t_1-t_2-t_3)^{-1}(2-t_2-t_3)^{-1}(1-t_3)^{-1},\ \ t_1+t_2+t_3<3, t_2+t_3<2, t_3<1 \end{aligned}$$

이로부터 $Y=(Y_1,Y_2,Y_3)^t$의 결합적률생성함수는

$$\begin{aligned} M_Y(s_1,s_2,s_3) &= E(e^{s_1Y_1+s_2Y_2+s_3Y_3}) \\ &= E[e^{s_1X_1+s_2(X_2-X_1)+s_3(X_3-X_2)}] \\ &= E[e^{(s_1-s_2)X_1+(s_2-s_3)X_2+s_3X_3}] \\ &= 6(3-(s_1-s_2)-(s_2-s_3)-s_3)^{-1}(2-(s_2-s_3)-s_3)^{-1}(1-s_3)^{-1} \\ &= 6(3-s_1)^{-1}(2-s_2)^{-1}(1-s_3)^{-1},\ \ s_1<3, s_2<2, s_3<1 \end{aligned}$$

로서 $m_1(s_1)m_2(s_2)m_3(s_3)$꼴로 주어진다. 따라서Y_1,Y_2,Y_3는 서로 독립이다.

정리 2.5.10: 독립인 다차원 확률변수들의 성질(p.102)
(a)(독립인 확률변수들의 함수) 다차원 확률변수 $X_1,\cdots,X_n$이 서로 독립이면 각각의 함수인 $g_1(X_1),\ \cdots,g_n(X_n)$도 서로 독립이다.
(b)(독립인 확률변수들의 곱의 기댓값[24]) 다차원 확률변수 $X_1,\cdots,X_n$이 서로 독립이면, 실수 값 함수 $g_1,\cdots,g_n$에 대하여

$$E[g_1(X_1)\cdots g(X_n)]=E[g_1(X_1)]\cdots E[g_n(X_n)]$$

(c)(독립성과 상관관계[25]) 다차원 확률변수 $X_1,\cdots,X_n$이 서로 독립이면,

$$Cov(X_i,X_j)=0\ \ (i\neq j)(i,j=1,\cdots,n)$$

정리 2.5.11: 서로 독립인 확률변수의 합(p.103)
(a)(합의 분산행렬[26])다차원 확률변수 $X_1,\cdots,X_n$이 서로 독립이면

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

(b)(합의 적률생성함수[27])다차원 확률변수 $X_1,\cdots,X_n$이 서로 독립이면

$$mgf_{X_1+\cdots+X_n}(t)=mgf_{X_1}(t)\cdots mgf_{X_n}(t)$$

[증명](a) 공분산행렬의 선형성을 뜻하는 정리 2.5.3의 (d)로부터

$$Var(X_1+\cdots+X_n)=Cov(\sum_{i=1}^n X_i,\sum_{j=1}^n X_j)=\sum_{i=1}^n\sum_{j=1}^n Cov(X_i,X_j)$$

따라서 정리 2.5.10의 독립성과 상관관계로부터 성질 (a)가 성립하는 것을 알 수 있다.
(b) 적률생성함수의 정의와 정리 2.5.10의 성질 (b)로부터

$$mgf_{X_1+\cdots+X_n}(t)=E[e^{t^t(X_1+\cdots+X_n)}]=E(e^{t^tX_1})\cdots E(e^{t^tX_n})=mgf_{X_1}(t)\cdots mgf_{X_n}(t)$$

24) 이 성질은 $g_1(X_1),\cdots,g_n(X_n)$의 기댓값이 실수로 정의된다는 것을 전제로 한다.
25) 이 성질은 두 확률변수 사이의 공분산 행렬이 실수 행렬로 정의된다는 것을 전제로 한다.
26) 이 성질은 각 확률변수의 분산행렬이 실수 행렬로 정의된다는 것을 전제로 한다.
27) 이 성질은 각 확률변수의 적률생성함수가 원점 근방의 열린 집합에서 정의될 수 있는 것을 전제로 한다.