4장. 표본 분포

4.3 순서통계량의 분포

예 4.3.1

표준지수분포 Exp(1)에서의 랜덤표본 X_1, X_2, X_3을 크기 순서로 늘어놓은 순서통계량을 $X_{(1)} < X_{(2)} < X_{(3)}$ 이라고 할 때 $Y = (X_{(1)}, X_{(2)}, X_{(3)})^t$ 의 결합확률밀도함수를 구하여라.

[풀이] 순서통계량을 나타내는 함수를 $u(X_1, X_2, X_3) = (X_{(1)}, X_{(2)}, X_{(3)})^t$라고 하면, 함수 u는

$$\mathcal{X} = \left\{(x_1, x_2, x_3)^t : x_i > 0\ (i = 1,2,3), x_1 \neq x_2, x_2 \neq x_3, x_3 \neq x_1\right\}$$

에서

$$\mathcal{Y} = \left\{(y_1, y_2, y_3)^t : 0 < y_1 < y_2 < y_3\right\}$$

로의 **3!대일 함수로서 다대일 변환**에 관한 정리 4.1.2의 조건을 만족하는 것을 알 수 있다.

구체적으로, $\{1,2,3\}$에서의 치환(置換 permutation)을 π로 나타내고

$$\mathcal{X}_\pi = \left\{(x_1, x_2, x_3)^t : 0 < x_{\pi 1} < x_{\pi 2} < x_{\pi 3}\right\}$$

라고 하면 **함수 u에서 정의역을 $\mathcal{X}_\pi$ 로 제한한 함수**

$$u^\pi(x_1, x_2, x_3) = (x_{\pi 1}, x_{\pi 2}, x_{\pi 3})^t,\ x \in \mathcal{X}_\pi$$

는 각각 $\mathcal{X}_\pi$에서 $\mathcal{Y}$ 로의 일대일 함수로서 미분가능하며

$$J_{u^\pi}(x) = \det\left(\frac{\partial x_{\pi j}}{\partial x_i}\right) = \pm 1 \quad \forall x \in \mathcal{X}_\pi$$

임을 알 수 있다. 예를 들어 $(\pi 1, \pi 2, \pi 3) = (3,2,1)$인 경우에

$$J_{u^\pi}(x) = \det\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = -1$$

따라서 정리 4.1.2의 공식으로부터 $Y = (X_{(1)}, X_{(2)}, X_{(3)})^t$ 의 결합확률밀도함수는 다음과 같이 주어진다.

$$\begin{aligned} pdf_Y(y) &= \sum_\pi pdf_X\left((u^\pi)^{-1}(y)\right)\left|J_{(u^\pi)^{-1}}(y)\right|,\ y \in \mathcal{Y} \\ &= \sum_\pi pdf_X(y_{\pi^{-1}1}, y_{\pi^{-1}2}, y_{\pi^{-1}3})\left|\pm 1\right|, y \in \mathcal{Y} \end{aligned}$$

한편 $X = (X_1, X_2, X_3)^t$의 결합확률밀도함수는

$$pdf_X(x_1, x_2, x_3) = e^{-(x_1 + x_2 + x_3)} I_{(x_1 > 0, x_2 > 0, x_3 > 0)}$$

이고 치환 π의 개수는 3!=6 개이므로

$$\begin{aligned} pdf_Y(y_1, y_2, y_3) &= \sum_\pi e^{-(y_{\pi^{-1}1} + y_{\pi^{-1}2} + y_{\pi^{-1}3})}, y \in \mathcal{Y} \\ &= 6e^{-(y_1 + y_2 + y_3)} I_{(0 < y_1 < y_2 < y_3)} \end{aligned}$$

정리 4.3.1: 순서통계량의 결합확률밀도함수

모집단 분포가 연속형이고 그 확률밀도함수가 $f(x)$일 때, 랜덤표본 $X_1,X_2,\cdots,X_n$을 크기 순서로 늘어놓은 순서통계량을 $X_{(1)} < X_{(2)} < \cdots < X_{(n)}$ 이라고 하면 $Y=(X_{(1)},X_{(2)}\cdots,X_{(n)})^t$ 의 결합확률밀도함수는 다음과 같이 주어진다.

$$pdf_Y(y_1,\cdots,y_n) = n!f(y_1)\cdots f(y_n)I_{(y_1<\cdots<y_n)}$$

[증명] 순서통계량을 나타내는 함수를 $u(X_1,\cdots,X_n)=(X_{(1)},\cdots,X_{(n)})^t$이라고 하면, 함수 u는

$$\mathcal{X} = \left\{(x_1,\cdots,x_n)^t : f(x_i)>0\ (i=1,\cdots,n),\ x_1,\cdots,x_n\text{은 서로 다른 실수}\right\}$$

에서

$$\mathcal{Y} = \left\{(y_1,\cdots,y_n)^t : f(y_i)>0\,(i=1,\cdots,n),\ y_1<\cdots<y_n\right\}$$

로의 $n!$**대일 함수로서 다대일 변환**에 관한 정리 4.1.2의 조건을 만족하는 것을 알 수 있다.

구체적으로, $\{1,\cdots,n\}$에서의 치환(置換 permutation)을 π로 나타내고

$$\mathcal{X}_\pi = \left\{(x_1,\cdots,x_n)^t : f(x_i)>0\ (i=1,\cdots,n),\ x_{\pi 1}<\cdots<x_{\pi n}\right\}$$

라고 하면 **함수 u에서 정의역을 $\mathcal{X}_\pi$로 제한한 함수**

$$u^\pi(x_1,\cdots,x_n) = (x_{\pi 1},\cdots,x_{\pi n})^t,\ x\in\mathcal{X}_\pi$$

는 각각 $\mathcal{X}_\pi$에서 Y 로의 일대일 함수로서 미분가능하며 일차편도함수의 행렬 $\left(\dfrac{\partial x_{\pi j}}{\partial x_i}\right)$은 단위행렬의 행과 열을 치환해놓은 행렬인 것을 알 수 있다. 따라서

$$J_{u^\pi}(x) = \det\left(\frac{\partial x_{\pi j}}{\partial x_i}\right) = \pm 1\quad \forall x\in\mathcal{X}_\pi$$

이고 정리 4.1.2의 공식으로부터 $Y=(X_{(1)},\cdots,X_{(n)})^t$ 의 결합확률밀도함수는

$$\begin{aligned} pdf_Y(y) &= \sum_\pi pdf_X((u^\pi)^{-1}(y))|J_{(u^\pi)^{-1}}(y)|,\ y\in\mathcal{Y} \\ &= \sum_\pi pdf_X(y_{\pi^{-1}1},\cdots,y_{\pi^{-1}n})|\pm 1|, y\in\mathcal{Y} \end{aligned}$$

한편 $X=(X_1,\cdots,X_n)^t$의 결합확률밀도함수는

$$pdf_X(x_1,\cdots,x_n) = f(x_1)\cdots f(x_n)$$

이고 치환 π의 개수는 $n!$ 개이므로

$$\begin{aligned} pdf_Y(y_1,\cdots,y_n) &= \sum_\pi f(y_{\pi^{-1}1})\cdots f(y_{\pi^{-1}n}), y\in\mathcal{Y} \\ &= n!f(y_1)\cdots f(y_n)I_{(y_1<\cdots<y_n)} \end{aligned}$$

순서통계량의 주변확률밀도함수의 직관적 유도 과정:

$X_{(r)}$이 x 근방에 있을 사건 $(x<X_{(r)}\le x+|\triangle x|)$은 $X_1,\cdots,X_n$들 중에서 $(r-1)$개는 x이하이고 1개는 x와 $x+|\triangle x|$사이에 그리고 나머지 $(n-r)$개는 $x+|\triangle x|$를 초과하는 사건으로 나타내어진다. 그런데 $X_1,\cdots,X_n$은 서로 독립이고 동일한 분포를 따르므로 이러한 사건의 확률은 다음과 같이 나타낼 수 있다.

$$P(x<X_{(r)}\le x+|\triangle x|) \approx c_r[P(X_1\le x)]^{r-1}P(x<X_1\le x+|\triangle x|)[P(X_1>x+|\triangle x|)]^{n-r}$$

여기에서 c_r은 서로 다른 n개 중에서 $(r-1)$개, 1개, $(n-r)$개를 뽑아서 각각 세 개의 구간 $(-\infty,x], (x,x+|\triangle x|], (x+|\triangle x|,+\infty)$에 넣는 방법의 수를 나타내는 상수이다. 즉

$$c_r = \frac{n!}{(r-1)!1!(n-r)!}$$

따라서 모집단의 확률밀도함수와 누적분포함수를 각각 f와 F라고 하면

$$pdf_{X_{(r)}}(x)|\triangle x| \approx c_r[F(x)]^{r-1}f(x)|\triangle x|[1-F(x)]^{n-r}$$

정리 4.3.2: 순서통계량의 주변확률밀도함수

모집단 분포가 연속형이고 그 확률밀도함수가 $f(x)$이고 누적분포함수가 $\mathrm{F}(x)$일 때, 랜덤표본 $\mathrm{X}_1,\cdots,\mathrm{X}_n$을 크기 순서로 늘어놓은 순서통계량을 $\mathrm{X}_{(1)}<\cdots<\mathrm{X}_{(n)}$이라고 하면 이들의 일차원 또는 이차원 주변확률밀도함수가 다음과 같이 주어진다.

(a) ($\mathrm{X}_{(r)}$의 확률밀도함수) $(1 \le r \le n)$

$$pdf_{\mathrm{X}_{(r)}}(x)=\frac{n!}{(r-1)!(n-r)!}[\mathrm{F}(x)]^{r-1}f(x)[1-\mathrm{F}(x)]^{n-r}$$

(b)($(\mathrm{X}_{(r)},\mathrm{X}_{(s)})^t$의 확률밀도함수) $(1 \le r < s \le n)$

$$pdf_{\mathrm{X}_{(r)},\mathrm{X}_{(s)}}(x,y)=\frac{n!}{(r-1)!(s-1-r)!(n-s)!}[\mathrm{F}(x)]^{r-1}f(x)[\mathrm{F}(y)-\mathrm{F}(x)]^{s-1-r}f(y)[1-\mathrm{F}(y)]^{n-s}$$

[증명]정리 4.3.1에 주어진 $(\mathrm{X}_{(1)},\cdots,\mathrm{X}_{(n)})^t$의 결합확률밀도함수를 $f_{1,\cdots,n}(y_1,\cdots,y_n)$으로 나타내고, $\mathrm{X}_{(r)}$과 $(\mathrm{X}_{(r)},\mathrm{X}_{(s)})^t$의 주변확률밀도함수를 각각 $f_r(x), f_{r,s}(x,y)$로 나타내면

$$f_r(x)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f_{1,\cdots,n}(y_1,\cdots,x,\cdots,y_n)dy_1\cdots dy_{r-1}dy_{r+1}\cdots dy_n$$

$$f_{r,s}(x,y)=\int_{-\infty}^{+\infty}\cdots\int_{-\infty}^{+\infty}f_{1,\cdots,n}(y_1,\cdots,x,\cdots,y,\cdots y_n)dy_1\cdots dy_{r-1}dy_{r+1}\cdots dy_{s-1}dy_{s+1}\cdots dy_n$$

이러한 적분에서 피적분함수의 적분 변수들에 대한 대칭성(symmetry)을 이용....

예 4.3.2 균등분포에서의 순서통계량:

균등분포 $\mathrm{U}(0,1)$에서의 랜덤표본 n개에 기초한 순서통계량을 $\mathrm{X}_{(1)}<\cdots<\mathrm{X}_{(n)}$이라고 할 때
다음을 밝혀라.

(a) $\mathrm{X}_{(r)} \sim \mathrm{Beta}(r, n-r+1)$, 즉 $\mathrm{X}_{(\mathrm{r})}$의 확률밀도함수는

$$f_r(x)=\frac{n!}{(r-1)!(n-r)!}x^{r-1}(1-x)^{n-r}\mathrm{I}_{(0,1)}(x)$$

(b) $\mathrm{Z}=(\mathrm{X}_{(1)},\mathrm{X}_{(2)}-\mathrm{X}_{(1)},\cdots,\mathrm{X}_{(n)}-\mathrm{X}_{(n-1)})^t \sim \mathrm{Dirichlet}(1,\cdots,1)$, 즉 Z의 확률밀도함수는

$$pdf_{\mathrm{Z}}(z_1,\cdots,z_n)=\frac{\Gamma(n+1)}{\Gamma(1)\cdots\Gamma(1)}I_{(z_1>0,\cdots,z_n>0,z_1+\cdots z_n<1)}$$

[풀이] (a)는 정리 4.3.2로부터 명백하다. 또한 정리 4.3.1로부터 $\mathrm{Y}=(\mathrm{X}_{(1)},\cdots,\mathrm{X}_{(n)})^t$의 결합확률밀도함수는

$$pdf_{\mathrm{Y}}(y_1,\cdots,y_n)=n!\mathrm{I}_{(0<y_1<\cdots<y_n<1)}$$

한편 $\mathrm{Z}=(\mathrm{X}_{(1)},\mathrm{X}_{(2)}-\mathrm{X}_{(1)},\cdots,\mathrm{X}_{(n)}-\mathrm{X}_{(n-1)})^t$와 $\mathrm{Y}=(\mathrm{X}_{(1)},\cdots,\mathrm{X}_{(n)})^t$에 대하여

$$\mathrm{Z}=(\mathrm{Z}_1,\cdots,\mathrm{Z}_n)^t=u((\mathrm{Y}_1,\cdots,\mathrm{Y}_n)^t)=u(\mathrm{Y})$$

라고 하면 다음과 같이 함수 u와 그 정의역 $\mathcal{Y}$, 치역 $\mathcal{Z}$,역함수 u^{-1}를 생각할 수 있다.

$$u:\begin{cases} z_1=y_1 \\ z_2=y_2-y_1 \\ \quad\vdots \\ z_n=y_n-y_{n-1}\end{cases} \qquad u^{-1}:\begin{cases} y_1=z_1 \\ y_2=z_1+z_2 \\ \quad\vdots \\ y_n=z_1+z_2+\cdots+z_n\end{cases}$$

$$\mathcal{Y}=\left\{(y_1,\cdots,y_n)^t : 0<y_1<\cdots<y_n<1\right\}$$

$$\mathcal{Z}=\left\{(z_1,\cdots,z_n)^t : z_i>0\,(i=1,\cdots,n), z_1+\cdots+z_n<1\right\}$$

<u>함수 u는 $\mathcal{Y}$에서 $\mathcal{Z}$로의 일대일 함수로서 정리 4.1.1의 조건을 만족시키고 그 역함수 u^{-1}의 야코비안 값은 1이므로</u> $\mathrm{Z}=(\mathrm{X}_{(1)},\mathrm{X}_{(2)}-\mathrm{X}_{(1)},\cdots,\mathrm{X}_{(n)}-\mathrm{X}_{(n-1)})^t$의 결합확률밀도함수는

$$pdf_{\mathrm{Z}}(z_1,\cdots,z_n)=n!\,\mathrm{I}_{(z_1>0,\cdots,z_n>0,z_1+\cdots z_n<1)}$$

예 4.3.3 지수분포에서의 순서통계량:

지수분포 Exp(1)에서의 랜덤표본 n개에 기초한 순서통계량 $X_{(1)} < \cdots < X_{(n)}$ 에 대하여

$$Z_1 = nX_{(1)}, Z_2 = (n-1)(X_{(2)} - X_{(1)}), \cdots, Z_n = X_{(n)} - X_{(n-1)}$$

이라고 할 때 $Z_1, Z_2, \cdots, Z_n$은 서로 독립이고 각각 지수분포 Exp(1)를 따르는 것을 밝혀라.

[풀이] $Y = (Y_1, Y_2, \cdots, Y_n) = (X_{(1)}, X_{(2)} \cdots, X_{(n)})^t$ 이라고 하면 정리 4.3.1로부터

$$pdf_Y(y_1, \cdots, y_n) = n!\, e^{-(y_1 + \cdots + y_n)} I_{(0 < y_1 < \cdots < y_n)}$$

한편 $Z = (nX_{(1)}, (n-1)(X_{(2)} - X_{(1)}), \cdots, X_{(n)} - X_{(n-1)})^t$와 $Y = (X_{(1)}, X_{(2)}, \cdots, X_{(n)})^t$에 대하여

$$Z = (Z_1, \cdots, Z_n)^t = u((Y_1, \cdots, Y_n)^t) = u(Y)$$

라고 하면 다음과 같이 함수 u와 그 정의역 $\mathcal{Y}$, 치역 $\mathcal{Z}$,역함수 u^{-1}를 생각할 수 있다.

$$u : \begin{cases} z_1 = ny_1 \\ z_2 = (n-1)(y_2 - y_1) \\ \quad \vdots \\ z_n = y_n - y_{n-1} \end{cases} \qquad u^{-1} : \begin{cases} y_1 = \dfrac{1}{n} z_1 \\ y_2 = \dfrac{1}{n} z_1 + \dfrac{1}{n-1} z_2 \\ \quad \vdots \\ y_n = \dfrac{1}{n} z_1 + \dfrac{1}{n-1} z_2 + \cdots + \dfrac{1}{1} z_n \end{cases}$$

$$\mathcal{Y} = \left\{ (y_1, \cdots, y_n)^t : 0 < y_1 < \cdots < y_n \right\}$$

$$\mathcal{Z} = \left\{ (z_1, \cdots, z_n)^t : z_1 > 0, \cdots, z_n > 0 \right\}$$

<u>함수 u는 $\mathcal{Y}$ 에서 $\mathcal{Z}$ 로의 일대일 함수로서 정리 4.1.1의 조건을 만족시키고 그 역함수 u^{-1} 의 야코비안 값은 $1/n!$, $y_1 + \cdots + y_n = z_1 + \cdots + z_n$</u>이므로 Z의 결합확률밀도함수는

$$pdf_Z(z_1, \cdots, z_n) = n!\, e^{-(z_1 + \cdots + z_n)} I_{(z_1 > 0, \cdots, z_n > 0)} \frac{1}{n!} = \prod_{i=1}^{n} \left\{ e^{-z_i} I_{(0, \infty)}(z_i) \right\}$$

지수분포에서의 순서통계량 분포의 대의적 정의:

지수분포 Exp(1)에서의 랜덤표본 n개에 기초한 순서통계량을 $X_{(1)} < \cdots < X_{(n)}$ 이라고 하면

$$(X_{(r)})_{1 \le r \le n} \overset{d}{\equiv} \left(\frac{1}{n} Z_1 + \cdots + \frac{1}{n-r+1} Z_r \right)_{1 \le r \le n}, \quad Z_r \overset{iid}{\sim} \text{Exp}(1)\ (r = 1, \cdots, n)$$

정리 4.3.3: 확률적분변환(確率積分變換 probability integral transformation)[1)]

확률변수 X가 연속형이고 그 누적분포함수 F가 순증가함수 즉

$$x_1 < x_2 \text{이면 } \mathrm{F}(x_1) < \mathrm{F}(x_2)$$

일 때 그 역함수를 F^{-1}라고 하면 다음이 성립한다.

(a) $\mathrm{F}(\mathrm{X}) \sim \mathrm{U}(0,1)$, 즉 $\mathrm{F}(\mathrm{X})$는 균등분포 $\mathrm{U}(0,1)$을 따른다.

(b) 균등분포 $\mathrm{U}(0,1)$을 따르는 확률변수 U에 대하여

$$\mathrm{F}^{-1}(\mathrm{U}) \overset{d}{\equiv} \mathrm{X}$$

즉 $\mathrm{F}^{-1}(\mathrm{U})$의 누적분포함수는 F로서 $\mathrm{F}^{-1}(\mathrm{U})$와 X는 같은 분포를 갖는다.

[증명] 주어진 조건하에서 역함수 F^{-1}도 순증가하고 다음이 성립하는 것을 알 수 있다.

$$\mathrm{F}(x) \le u \Leftrightarrow x \le \mathrm{F}^{-1}(u),\quad \mathrm{F}^{-1}(u) \le x \Leftrightarrow u \le \mathrm{F}(x),\quad \mathrm{F}(\mathrm{F}^{-1}(u)) = u$$

(a) $$\mathrm{P}(\mathrm{F}(\mathrm{X}) \le u) = \mathrm{P}(\mathrm{X} \le \mathrm{F}^{-1}(u)) = \mathrm{F}(\mathrm{F}^{-1}(u)) = u \quad \forall u : 0 \le u \le 1$$

$$\therefore\ \mathrm{F}(\mathrm{X}) \sim \mathrm{U}(0,1)$$

(b) $$\mathrm{P}(\mathrm{F}^{-1}(\mathrm{U}) \le x) = \mathrm{P}(\mathrm{U} \le \mathrm{F}(x)) = \mathrm{F}(x) \quad \forall x : -\infty < x < +\infty$$

$$\therefore\ \mathrm{F}^{-1}(\mathrm{U}) \overset{d}{\equiv} \mathrm{X}$$

예 4.3.4 균등분포와 지수분포에서의 순서통계량:

균등분포 $\mathrm{U}(0,1)$에서의 랜덤표본 n개에 기초한 순서통계량을 $\mathrm{U}_{(1)} < \cdots < \mathrm{U}_{(n)}$이라고 하고지수분포 $\mathrm{Exp}(1)$에서의 랜덤표본 n개에 기초한 순서통계량을 $\mathrm{Y}_{(1)} < \cdots < \mathrm{Y}_{(n)}$이라고 하면

$$\left(\mathrm{U}_{(r)}\right)_{1 \le r \le n} \overset{d}{\equiv} \left(1 - e^{-\mathrm{Y}_{(r)}}\right)_{1 \le r \le n}$$

[풀이] 지수분포 $\mathrm{Exp}(1)$에서의 랜덤표본을 $\mathrm{Y}_1, \mathrm{Y}_2, \cdots, \mathrm{Y}_n$이라고 할 때, 지수분포의 누적분포함수가

$$\mathrm{G}(y) = \begin{cases} 1 - e^{-y}, & y \ge 0 \\ 0, & y < 0 \end{cases}$$

이므로 정리 4.3.3으로부터

$$\mathrm{G}(\mathrm{Y}_i) \overset{iid}{\sim} \mathrm{U}(0,1)$$

또한 누적분포함수는 증가함수이므로

$$\left(\mathrm{G}(\mathrm{Y}_{(r)})\right)_{1 \le r \le n} \overset{d}{\equiv} \left(\mathrm{U}_{(r)}\right)_{1 \le r \le n}$$

---

1) 누적분포함수가 순증가는 아닌 증가함수인 경우에는 F의 역변환을

$$\mathrm{F}^{-1}(u) = inf\{x : \mathrm{F}(x) \ge u\}\ (0 \le u \le 1)$$

로 정의하고, 이 역변환에 대하여 정리의 (b)가 일반적인 확률변수에 대하여도 성립한다.

정리 4.3.4: 순서통계량 분포의 대의적 정의

모집단 분포가 연속형이고 그 누적분포함수가 $\mathrm{F}(x)$일 때 랜덤표본 n개에 기초한 순서통계량을 $\mathrm{X}_{(1)} < \cdots < \mathrm{X}_{(n)}$이라고 하면 $h(y) = \mathrm{F}^{-1}(1-e^{-y})\mathrm{I}_{(0,+\infty)}(y)$로 정의된 함수 h에 대하여

$$(\mathrm{X}_{(r)})_{1 \le r \le n} \overset{d}{\equiv} \left(h(\frac{1}{n}\mathrm{Z}_1 + \cdots + \frac{1}{n-r+1}\mathrm{Z}_r)\right)_{1 \le r \le n}, \quad \mathrm{Z}_r \overset{iid}{\sim} \mathrm{Exp}(1)\ (r = 1, \cdots, n)$$

[증명] 균등분포 $\mathrm{U}(0,1)$에서의 랜덤표본 n개에 기초한 순서통계량을 $\mathrm{U}_{(1)} < \cdots < \mathrm{U}_{(n)}$이라고 하면 정리 4.3.3으로부터

$$(\mathrm{X}_{(r)})_{1 \le r \le n} \overset{d}{\equiv} \left(\mathrm{F}^{-1}(\mathrm{U}_{(r)})\right)_{1 \le r \le n}$$

한편 지수분포 $\mathrm{Exp}(1)$에서의 랜덤표본 n개에 기초한 순서통계량을 $\mathrm{Y}_{(1)} < \cdots < \mathrm{Y}_{(n)}$이라고 하면 예 4.3.4로부터

$$(\mathrm{U}_{(r)})_{1 \le r \le n} \overset{d}{\equiv} \left(1 - e^{-\mathrm{Y}_{(r)}}\right)_{1 \le r \le n}$$

$$\therefore\ (\mathrm{X}_{(r)})_{1 \le r \le n} \overset{d}{\equiv} \left(h(\mathrm{Y}_{(r)})\right)_{1 \le r \le n}$$

또한 예 4.3.3으로부터

$$(\mathrm{Y}_{(r)})_{1 \le r \le n} \overset{d}{\equiv} \left(\frac{1}{n}\mathrm{Z}_1 + \cdots + \frac{1}{n-r+1}\mathrm{Z}_r\right)_{1 \le r \le n}, \quad \mathrm{Z}_r \overset{iid}{\sim} \mathrm{Exp}(1)\ (r = 1, \cdots, n)$$

따라서 정리에 주어진 대의적 정의를 얻게 된다.

4.4 다변량 정규분포(多變量 正規分布 multivariate normal distribution)

정리 4.4.1
표준정규분포 $\mathrm{N}(0,1)$을 따르고 서로 독립인 확률변수 $\mathrm{Z}_1,\mathrm{Z}_2,\cdots,\mathrm{Z}_n$과 상수의 행렬과 벡터

$$A=(a_{ij})_{\substack{1\le i\le n\\ 1\le j\le n}},\ \ \mu=(\mu_1,\mu_2,\cdots,\mu_n)^t$$

에 대하여

$$\begin{cases}\mathrm{X}_1=a_{11}\mathrm{Z}_1+\cdots+a_{1n}\mathrm{Z}_n+\mu_1\\ \quad\vdots\\ X_n=a_{n1}\mathrm{Z}_1+\cdots+a_{nn}\mathrm{Z}_n+\mu_n\end{cases}\quad \text{즉 } \mathrm{X}=A\mathrm{Z}+\mu,\ \ \mathrm{X}=(\mathrm{X}_1,\cdots,\mathrm{X}_n)^t, \mathrm{Z}=(\mathrm{Z}_1,\cdots,\mathrm{Z}_n)^t$$

이라고 할 때, 다음이 성립한다.
(a) 행렬 A가 정칙행렬(정칙행렬 nonsingular matrix)이면 $\mathrm{X}=A\mathrm{Z}+\mu$의 확률밀도함수는

$$pdf_{\mathrm{X}}(x)=(\det(2\pi\Sigma))^{-1/2}\exp\{-(x-\mu)^t\Sigma^{-1}(x-\mu)/2\},\ x\in R^n\quad(\Sigma=AA^t)$$

(b) $\mathrm{X}=A\mathrm{Z}+\mu$의 적률생성함수는

$$mgf_{\mathrm{X}}(t)=\exp(\mu^t t+\frac{1}{2}t^t\Sigma t),\ \ t\in R^n$$

[증명] 정리 3.6.1로부터 $\mathrm{Z}=(\mathrm{Z}_1,\mathrm{Z}_2,\cdots,\mathrm{Z}_n)^t$의 확률밀도함수와 적률생성함수는 각각

$$pdf_{\mathrm{Z}}(z)=(2\pi)^{-\frac{n}{2}}e^{-\frac{1}{2}(z_1^2+\cdots+z_n^2)}=(2\pi)^{-\frac{n}{2}}e^{-\frac{1}{2}z^tz},\ \ z\in R^n$$

$$mgf_{\mathrm{Z}}(s)=e^{\frac{1}{2}(s_1^2+\cdots+s_n^2)}=e^{\frac{1}{2}s^ts},\ \ s\in R^n$$

(a) $\mathrm{X}=u(\mathrm{Z})=A\mathrm{Z}+\mu$에 대하여 정리 4.1.1 적용

$$pdf_{\mathrm{X}}(x)=pdf_{\mathrm{Z}}(u^{-1}(x))|J_{u^{-1}}(x)|=(2\pi)^{-n/2}\exp\left\{-\frac{1}{2}(A^{-1}(x-\mu))^tA^{-1}(x-\mu)\right\}(\det(A))^{-1}$$

$$pdf_{\mathrm{X}}(x)=(\det(2\pi\Sigma))^{-1/2}\exp\left\{-\frac{1}{2}(x-\mu)^t\Sigma^{-1}(x-\mu)\right\}\Sigma=AA^t$$

(b) $$mgf_{\mathrm{X}}(t)=\mathrm{E}[\exp\{t^t(A\mathrm{Z}+\mu)\}]=mgf_{\mathrm{Z}}(A^tt)\exp(t^t\mu)$$

$$\therefore\ mgf_X(t)=\exp\left\{\mu^tt+\frac{1}{2}(A^tt)^t(A^tt)\right\}=\exp(\mu^tt+\frac{1}{2}t^t\Sigma t)$$

정리 4.4.2
(a) 정리 4.4.1에서 $\mathrm{X}=A\mathrm{Z}+\mu$의 분포를 $\mathrm{X}\sim\mathrm{N}(\mu,\Sigma)$라고 나타내면

$$\mathrm{E}(\mathrm{X})=\mu,\ \ \mathrm{Var}(\mathrm{X})=\Sigma$$

(b) 분산행렬 Σ에 대하여 $AA^t=\Sigma, A=A^t$인 행렬 A를 $\Sigma^{1/2}$이라고 하면[2)]

$$\mathrm{X}\sim\mathrm{N}_n(\mu,\Sigma)\Leftrightarrow \mathrm{X}\overset{d}{\equiv}\Sigma^{1/2}\mathrm{Z}+\mu,\ \ \mathrm{Z}\sim\mathrm{N}_n(0,\mathrm{I})$$

[증명] (a) $\mathrm{N}(\mu,\Sigma)$의 누율생성함수 $\mathrm{C}(t)=\log mgf_{\mathrm{X}}(t)=\mu^tt+\frac{1}{2}t^t\Sigma t$ 의 미분

(b) $$mgf_{\Sigma^{1/2}\mathrm{Z}+\mu}(t)=\exp(\mu^tt+\frac{1}{2}t^t\Sigma t),\ \ t\in R^n$$

---

2) 정리 2.5.4로부터 분산행렬 Σ는 음이 아닌 정부호의 행렬이고, 부록 II에서 이러한 행렬에 대하여 $\Sigma^{1/2}\Sigma^{1/2}=\Sigma$인 대칭행렬 $\Sigma^{1/2}$이 존재하는 것을 밝히고 있다.

다변량 정규분포의 정의:

$Z \sim N_n(0,I) \Leftrightarrow Z=(Z_1,\cdots,Z_n)^t,\ Z_i \overset{iid}{\sim} N(0,1)\ (i=1,\cdots,n)$

$X \sim N_n(\mu,\Sigma)$ $\Leftrightarrow$(1)[3] $X \overset{d}{\equiv} AZ+\mu,\ Z \sim N(0,I),\ AA^t=\Sigma$

$\Leftrightarrow$(2) $X \overset{d}{\equiv} \Sigma^{1/2}Z+\mu,\ Z \sim N_n(0,I),\ \Sigma^{1/2}\Sigma^{1/2}=\Sigma,\ \Sigma^{1/2}=(\Sigma^{1/2})^t$

$\Leftrightarrow$(3) $mgf_X(t)=\exp(\mu^t t+\frac{1}{2}t^t\Sigma t),\ t\in R^n$

$\Leftrightarrow$(4) (분산행렬 Σ 가 정칙행렬인 경우)

$$pdf_X(x)=(\det(2\pi\Sigma))^{-1/2}\exp\left\{-\frac{1}{2}(x-\mu)^t\Sigma^{-1}(x-\mu)\right\},\ x\in R^n$$

예 4.4.1 이변량 정규분포 $N(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix})$ 또는 $N(\mu_1,\mu_2;\sigma_1^2,\sigma_2^2,\rho)$:

분산행렬 Σ에 대하여 $\det(\Sigma)=\sigma_1^2\sigma_2^2(1-\rho^2)$이므로, $\sigma_1>0,\sigma_2>0,\ -1<\rho<1$이면

$$\Sigma^{-1}=\frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}\begin{pmatrix}\sigma_2^2 & -\rho\sigma_1\sigma_2\\ -\rho\sigma_1\sigma_2 & \sigma_1^2\end{pmatrix}$$

$$pdf(x_1,x_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-\frac{1}{2}Q}$$

$$Q=\frac{1}{1-\rho^2}\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right\}$$

한편 $\sigma_1>0,\sigma_2>0,\rho=\pm1$ 인 경우에는 분산행렬의 역행렬이 존재하지 않고 정리 2.2.3으로부터 다음과 같이 두 변수 사이에 선형 관계가 성립하는 것을 알 수 있다.

$$\rho=1 \Leftrightarrow P(\frac{X_2-\mu_2}{\sigma_2}=\frac{X_1-\mu_1}{\sigma_1})=1, \qquad \rho=-1 \Leftrightarrow P(\frac{X_2-\mu_2}{\sigma_2}=-\frac{X_1-\mu_1}{\sigma_1})=1$$

정리 4.4.3: 다변량 정규분포의 성질

(a) $X \sim N(\mu,\Sigma)$ 이면 상수의 행렬과 벡터 A와 b에 대하여

$$AX+b \sim N(A\mu+b, A\Sigma A^t)$$

(b) $\begin{pmatrix}X_1\\ X_2\end{pmatrix} \sim N(\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix},\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix})$ 일 때, $Cov(X_1,X_2)=\Sigma_{12}=0$이면 X_1과 X_2는 서로 독립이다.

(c) $X \sim N(\mu,\Sigma)$이고 A,B가 상수의 행렬일 때, $Cov(AX,BX)=A\Sigma B^t=0$이면 AX와 BX는 서로 독립이다.

[증명] (a) 다변량 정규분포의 정의 (2)와 (1) 적용

(b) 정의 (3)으로부터 주어지는 다음과 같은 X_1과 X_2의 결합적률생성함수이용

$$mgf_{X_1,X_2}(t_1,t_2)=\exp\left\{\begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix}^t\begin{pmatrix}t_1\\ t_2\end{pmatrix}+\frac{1}{2}\begin{pmatrix}t_1\\ t_2\end{pmatrix}^t\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}\begin{pmatrix}t_1\\ t_2\end{pmatrix}\right\}$$

(c) (a)와 (b) 이용

3) 정의 (1)에서 A가 $n\times n$정방행렬이어야만 하는 것은 아니며 $AA^t=\Sigma$인 $n\times m$ 행렬이어도 무방하다.

정리 4.4.4: 다변량 정규분포의 주변분포와 조건부분포

(a) (주변분포) $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$이면

$$X_1 \sim N(\mu_1, \Sigma_{11})$$

(b) (조건부분포) $\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right)$이고 분산행렬이 정칙행렬이면

$$X_{2|X_1 = x_1} \sim N(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1), \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$$

[증명] (a) $X_1 = (I, 0)\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ 와 정리 4.4.3의 (a) 이용

(b) 이 경우의 분산행렬과 같이 부분행렬로 분할되어 있는 정칙행렬에 대하여

$$\det\left[\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\right] = \det(\Sigma_{11})\det(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$$

이고 $\Sigma_{22 \cdot 1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$라고 하면

$$\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \Sigma_{11}^{-1} + \Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22 \cdot 1}^{-1}\Sigma_{21}\Sigma_{11}^{-1} & -\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22 \cdot 1}^{-1} \\ -\Sigma_{22 \cdot 1}^{-1}\Sigma_{21}\Sigma_{11}^{-1} & \Sigma_{22 \cdot 1}^{-1} \end{pmatrix}$$

임이 알려져 있다.[4] 이 사실을 이용하여 **조건부분포의 확률밀도함수**

$$pdf_{X_2|X_1 = x_1}(x_2|x_1) = pdf_{X_1, X_2}(x_1, x_2)/pdf_{X_1}(x_1)$$

를 계산하여 (b)를 밝힐 수 있다.

또는 (b)와

$$X_2 - \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1)_{|X_1 = x_1} \sim N(0, \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$$

가 동등한 것을 이용하여 다음과 같이 (b)를 밝힐 수도 있다.

$$\begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1) \end{pmatrix} = \begin{pmatrix} I & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 \end{pmatrix}$$

$$\begin{pmatrix} I & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}\begin{pmatrix} I & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I \end{pmatrix}^t = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{pmatrix}$$

따라서 정리 4.4.3의 (a)로부터

$$\begin{pmatrix} X_1 - \mu_1 \\ X_2 - \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1) \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \end{pmatrix}\right)$$

그러므로 정리 4.4.3의 (b)로부터 $X_1 - \mu_1$과 $X_2 - \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1)$은 서로 독립이고
$X_2 - \mu_2 - \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1) \sim N(0, \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})$
이다. 따라서 (b)의 조건부분포가 성립한다.

예 4.4.3 이변량 정규분포 $N(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2, \rho)$에서의 조건부분포:

$$X_{2|X_1 = x_1} \sim N(\mu_2 + \rho\sigma_2(x_1 - \mu_1)/\sigma_1, \sigma_2^2(1 - \rho^2))$$

---

4) 부분행렬로 분할된 행렬에 관한 이러한 사실은 부록 II에 증명되어 있다.

부록 II: 행렬 대수 (p.513~514)

분할된 행렬의 덧셈과 곱셈:

A_{ij}, B_{ij}가 행렬인 경우에

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}+B_{11} & A_{12}+B_{12} \\ A_{21}+B_{21} & A_{22}+B_{22} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11}+A_{12}B_{21} & A_{12}B_{12}+A_{12}B_{22} \\ A_{21}B_{11}+A_{22}B_{21} & A_{21}B_{12}+A_{22}B_{22} \end{pmatrix}$$

분할된 행렬의 행렬식:

(a) $\det\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \det(A)\det(B)$

(b) $\det\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \det(A)\det(B)$

(c) A_{11}^{-1}가 존재할 때, $\det\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \det(A_{11})\det(A_{22} - A_{21}A_{11}^{-1}A_{12})$

[증명] (a) 다음 등식과 행렬식의 귀납적 정의로부터 명백하다.

$$\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix}\begin{pmatrix} I & 0 \\ 0 & B \end{pmatrix}$$

(b) $\det(A) = 0$인 경우에는 (b)의 성립이 명백하고, $\det(A) \neq 0$인 경우에는 등식

$$\begin{pmatrix} A & C \\ 0 & B \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}\begin{pmatrix} I & A^{-1}C \\ 0 & I \end{pmatrix},\ \det\begin{pmatrix} I & D \\ 0 & I \end{pmatrix} = 1$$

로부터 (b)가 성립하는 것을 알 수 있다.

(c) 다음 등식과 (a), (b)로부터 (c)가 성립하는 것은 명백하다.

$$\begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix} = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{pmatrix}$$

분할된 행렬의 역행렬:

A_{11}^{-1}가 존재하고 $A_{22\cdot 1} = A_{22} - A_{21}A_{11}^{-1}A_{12}$의 역행렬도 존재할 때

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}A_{22\cdot 1}^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22\cdot 1}^{-1} \\ -A_{22\cdot 1}^{-1}A_{21}A_{11}^{-1} & A_{22\cdot 1}^{-1} \end{pmatrix}$$

[증명] 행렬식의 증명에서와 같이 다음 등식으로부터 역행렬을 구할 수 있다.

$$\begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}\begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix} = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{pmatrix}$$

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}^{-1}\begin{pmatrix} A_{11} & 0 \\ 0 & A_{22\cdot 1} \end{pmatrix}\begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix}^{-1}$$

$$\therefore \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} I & -A_{11}^{-1}A_{12} \\ 0 & I \end{pmatrix}\begin{pmatrix} A_{11}^{-1} & 0 \\ 0 & A_{22\cdot 1}^{-1} \end{pmatrix}\begin{pmatrix} I & 0 \\ -A_{21}A_{11}^{-1} & I \end{pmatrix}$$

대칭행렬의 대각화 정리:

원소가 실수인 $m \times m$ 행렬 A가 대칭행렬이면

$$Ax_i = \lambda_i x_i,\ x_i^t x_i = 1,\ x_i^t x_j = 0\,(i \neq j)\,(i,j = 1,2,\cdots,m)$$

를 만족시키는 실수의 고유값 λ_i와 실수 성분의 고유벡터 $x_i (i = 1,\cdots,m)$이 존재한다. 즉

$$A = P\,diag(\lambda_i)P^t,\ PP^t = P^tP = I$$

인 직교행렬 $P = (x_1, x_2, \cdots, x_m)$가 존재한다.

정리 4.4.5: 이차형식의 분포

(a) $X \sim N_k(\mu,\Sigma)$이고 분산행렬 Σ가 정칙행렬이면

$$(X-\mu)^t \Sigma^{-1}(X-\mu) \sim \chi^2(k)$$

(b) $Z \sim N(0,I)$일 때, $A^2 = A$이면 $Z^t AZ \sim \chi^2(r)$이고 자유도는 $r = trace(A)$이다.5)

[증명] (a) $\Sigma^{1/2}\Sigma^{1/2} = \Sigma$인 대칭행렬 $\Sigma^{1/2}$의 역행렬을 $\Sigma^{-1/2}$ 이라고 하면

$$\Sigma^{-1/2}(X-\mu) \sim N_k(0,I)$$

따라서 $Z = (Z_1,\cdots,Z_k)^t = \Sigma^{-1/2}(X-\mu)$라고 하면

$$(X-\mu)^t \Sigma^{-1}(X-\mu) = Z^t Z = Z_1^2 + \cdots + Z_k^2,\ Z_i \overset{iid}{\sim} N(0,1)(i=1,\cdots,k)$$

$$\therefore (X-\mu)^t \Sigma^{-1}(X-\mu) \sim \chi^2(k)$$

(b) 행렬 A는 원소가 실수인 대칭행렬이므로 다음과 같이 대각화가 가능하다.6)

$$A = P^t \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} P, \quad P^t P = PP^t = I$$

여기에서 $\lambda_1,\cdots,\lambda_n$은 $\det(A-\lambda I) = 0$을 만족하는 λ 값들로서, $A^2 = A$인 경우에는 $\lambda_i^2 = \lambda_i$ 이어야 하므로 λ_i는 0 또는 1 임을 알 수 있다. 이들 중 r 개만 1이고 나머지는 0이라고 하면 $\lambda_1 = \cdots = \lambda_r = 1,\ \lambda_{r+1} = \cdots = \lambda_n = 0$ 이라고 가정할 수 있다.

$$\therefore Z^t AZ = (PZ)^t \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} (PZ)$$

그러므로 $X = (X_1,\cdots,X_n)^t = PZ$라고 하면

$$Z^t AZ = X^t \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} X = X_1^2 + \cdots + X_r^2,\ X \sim N_n(0, PP^t) = N_n(0,I)$$

$$\therefore\ Z^t AZ = X_1^2 + \cdots + X_r^2,\ X_i \overset{iid}{\sim} N(0,1)(i=1,\cdots,r)$$

그러므로 카이제곱분포의 대의적 정의로부터

$$Z^t AZ \sim \chi^2(r)$$

---

5) 일반적으로 $x^t Ax = x^t(\frac{A+A^t}{2})x$ 이므로 이차형식 $Z^t AZ$에서 행렬 A는 언제나 대칭행렬인 것으로 가정되어 있으며, (b)에서의 조건 $A^2 = A$ 는 $Z^t AZ \sim \chi^2(r)$이기 위한 충분조건일 뿐만 아니라 필요조건인 것도 알려져 있다. 여기에서 $trace(A)$는 행렬 A의 모든 대각원소의 합이다.

6) 원소가 실수인 대칭행렬의 대각화에 대한 정리는 부록 II에 주어져 있으며, $\det(A-\lambda I) = 0$을 만족하는 λ 값은 행렬 A의 고유값(固有값 characteristic value)이라고 한다.

예 4.4.6 표본분산의 표본분포:

정규분포 $N(\mu,\sigma^2)$에서의 랜덤표본을 $X_1,X_2,\cdots,X_n$이라고 할 때, 표본평균 $\overline{X}$와 표본분산 $S^2=\sum_{i=1}^{n}(X_i-\overline{X})^2/(n-1)$은 서로 독립이고 $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$인 것을 정리 4.2.2로부터 이미 알고 있다. 이제 다변량 정규분포에 관한 정리들을 이용하여 이를 밝히는 방법을 알아보자. 이를 위하여 $X=(X_1,\cdots,X_n)^t$라고 하면

$$X \sim N_n(\mu 1, \sigma^2 I),\ \ 1=(1,\cdots,1)^t$$

$$\overline{X}=n^{-1}1^tX,\ (n-1)S^2=\sum_{i=1}^{n}(X_i-\overline{X})^2=X^t(I-n^{-1}11^t)X$$

한편 $I-n^{-1}11^t$은 대칭행렬이고 $1^t(I-n^{-1}11^t)=0$ 이므로

$$Cov(1^tX, (I-n^{-1}11^t)X)=1^t(\sigma^2 I)(I-n^{-1}11^t)^t=0$$

따라서 정리 4.4.3의 (c)로부터 $\overline{X}=n^{-1}1^tX$ 와 $(I-n^{-1}11^t)X$은 서로 독립이다. 그런데

$$(I-n^{-1}11^t)^2=I-n^{-1}11^t$$

이므로

$$X^t(I-n^{-1}11^t)X=[(I-n^{-1}11^t)X]^t[(I-n^{-1}11^t)X]$$

그러므로 $\overline{X}$와 $(n-1)S^2=X^t(I-n^{-1}11^t)X$은 서로 독립이다.

또한

$$X^t(I-n^{-1}11^t)X/\sigma^2=(X-\mu 1)^t(I-n^{-1}11^t)(X-\mu 1)/\sigma^2$$

이고

$$(X-\mu 1)/\sigma \sim N_n(0,I),\quad (I-n^{-1}11^t)^2=I-n^{-1}11^t,\quad trace(I-n^{-1}11^t)=n-1$$

이므로 정리 4.4.5의 (b)로부터

$$(n-1)S^2/\sigma^2=X^t(I-n^{-1}11^t)X/\sigma^2 \sim \chi^2(n-1)$$

선형회귀모형: 정규 오차항을 갖는 경우

$$\begin{cases} Y_i = x_{i0}\beta_0 + x_{i1}\beta_1 + \cdots + x_{ip}\beta_p + e_i \\ e_i \overset{iid}{\sim} N(0, \sigma^2),\ i = 1, \cdots, n \end{cases}$$

$$Y = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix},\ X = \begin{pmatrix} x_{10} & x_{11} & \cdots & x_{1p} \\ \vdots & \vdots & & \vdots \\ x_{n0} & x_{n1} & \cdots & x_{np} \end{pmatrix},\ \beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix},\ e = \begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix}$$

라고 하면 위의 모형을 다음과 같이 간략하게 나타낼 수 있다.

$$\begin{cases} Y = X\beta + e \\ e \sim N_n(0, \sigma^2 I),\ rank(X) = p+1 \end{cases}$$

표본회귀계수(표본회귀계수 sample regression coefficient):

$$\hat{\beta} = (X^t X)^{-1} X^t Y$$

평균오차제곱합(平均誤差제곱합 mean squared error sum of squares):

$$\widehat{\sigma^2} = (Y - X\hat{\beta})^t (Y - X\hat{\beta}) / (n-p-1) = \sum_{i=1}^{n} (Y_i - \widehat{\beta_0} x_{i0} - \widehat{\beta_1} x_{i1} - \cdots - \widehat{\beta_p} x_{ip})^2 / (n-p-1)$$

예 4.4.7 단순선형회귀모형에서 표본회귀계수:

선형회귀모형에서 $p = 1$이고 $x_0 = 1$ 경우, 즉

$$\begin{cases} Y_i = \beta_0 + x_{i1}\beta_1 + e_i \\ e_i \overset{iid}{\sim} N(0, \sigma^2),\ i = 1, \cdots, n \end{cases}$$

로 나타내어지는 모형을 단순(單純 simple)선형회귀모형이라고 한다. 이 경우에는

$$X = \begin{pmatrix} 1 & x_{11} \\ \vdots & \vdots \\ 1 & x_{n1} \end{pmatrix},\ X^t X = \begin{pmatrix} n & \sum_{i=1}^{n} x_{i1} \\ \sum_{i=1}^{n} x_{i1} & \sum_{i=1}^{n} x_{i1}^2 \end{pmatrix},\ X^t Y = \begin{pmatrix} \sum_{i=1}^{n} Y_i \\ \sum_{i=1}^{n} x_{i1} Y_i \end{pmatrix}$$

$$(X^t X)^{-1} = \frac{1}{n\sum_{i=1}^{n} x_{i1}^2 - (\sum_{i=1}^{n} x_{i1})^2} \begin{pmatrix} \sum_{i=1}^{n} x_{i1}^2 & -\sum_{i=1}^{n} x_{i1} \\ -\sum_{i=1}^{n} x_{i1} & n \end{pmatrix}$$

이므로 표본회귀계수 $\hat{\beta} = (\widehat{\beta_0}, \widehat{\beta_1})^t$는 다음과 같이 주어진다.

$$\widehat{\beta_1} = \frac{S_{xY}}{S_{xx}},\ S_{xY} = \sum_{i=1}^{n} (x_{i1} - \overline{x_1}) Y_i,\ S_{xx} = \sum_{i=1}^{n} (x_{i1} - \overline{x_1})^2,\ \overline{x_1} = \sum_{i=1}^{n} x_{i1}/n$$

$$\widehat{\beta_0} = \overline{Y} - \overline{x_1}\widehat{\beta_1},\quad \text{즉} \quad \widehat{\beta_0} + x_{i1}\widehat{\beta_1} = \overline{Y} + (x_{i1} - \overline{x_1})\widehat{\beta_1}$$

단순선형회귀모형에서 평균오차제곱합은

$$\widehat{\sigma^2} = \sum_{i=1}^{n} (Y_i - \widehat{\beta_0} - \widehat{\beta_1} x_{i1})^2 / (n-2) = \sum_{i=1}^{n} (Y_i - \overline{Y} - \widehat{\beta_1}(x_{i1} - \overline{x_1}))^2 / (n-2)$$

정리 4.4.6; 선형회귀모형에서 표본분포에 관한 기본 정리

선형회귀모형

$$\begin{cases} \mathrm{Y} = X\beta + e \\ e \sim \mathrm{N}_n(0, \sigma^2 \mathrm{I}) \end{cases}$$

에서 주어진 상수의 행렬 X는 $n\times(p+1)$행렬로서 계수가 $(p+1)$인 것으로 가정할 때 다음이 성립한다.

(a) $\hat{\beta} \sim \mathrm{N}_{\mathrm{p+1}}(\beta, \sigma^2(X^t X)^{-1})$

(b) 표본회귀계수 $\hat{\beta}$와 평균오차제곱합 $\widehat{\sigma^2}$은 서로 독립이다.

(c) $(n-p-1)\widehat{\sigma^2}/\sigma^2 \sim \chi^2(n-p-1)$

[증명] (a) $\mathrm{Y} \sim \mathrm{N}(X\beta, \sigma^2\mathrm{I})$와 정리 4.4.3의 (a)이용

(b) $\Pi = X(X^t X)^{-1} X^t$ 라고 하면 Π는 대칭행렬로서

$$\Pi^2 = X(X^t X)^{-1} X^t X (X^t X)^{-1} X^t = \Pi,\ \Pi X = X(X^t X)^{-1} X^t X = X$$

$$\mathrm{Y} - X\hat{\beta} = \mathrm{Y} - \Pi\mathrm{Y} = (\mathrm{I}-\Pi)\mathrm{Y}$$

$$\therefore\ \mathrm{Cov}((X^t X)^{-1} X^t \mathrm{Y}, (\mathrm{I}-\Pi)\mathrm{Y}) = (X^t X)^{-1} X^t (\sigma^2 \mathrm{I})(\mathrm{I}-\Pi)^t = 0$$

정리 4.4.3의 (c)로부터 $\hat{\beta} = (X^t X)^{-1} X^t \mathrm{Y}$와 $\mathrm{Y} - X\hat{\beta} = (\mathrm{I}-\Pi)\mathrm{Y}$는 서로 독립

따라서 $\hat{\beta}$와 $\widehat{\sigma^2} = (\mathrm{Y} - X\hat{\beta})^t (\mathrm{Y} - X\hat{\beta})/(n-p-1)$이 서로 독립

(c) (b)에서의 행렬 Π에 대하여

$$(\mathrm{I}-\Pi)^2 = \mathrm{I}-\Pi,\ (I-\Pi)X = 0$$

이므로

$$(\mathrm{Y} - X\hat{\beta})^t (\mathrm{Y} - X\hat{\beta}) = \mathrm{Y}^t (\mathrm{I}-\Pi)^t (\mathrm{I}-\Pi)\mathrm{Y} = \mathrm{Y}^t (\mathrm{I}-\Pi)\mathrm{Y} = (\mathrm{Y} - X\beta)^t (\mathrm{I}-\Pi)(\mathrm{Y} - X\beta)$$

$$\therefore\ (n-p-1)\widehat{\sigma^2}/\sigma^2 = (\mathrm{Y} - X\hat{\beta})^t (\mathrm{Y} - X\hat{\beta})/\sigma^2 = (\mathrm{Y} - X\beta)^t (\mathrm{I}-\Pi)(\mathrm{Y} - X\beta)/\sigma^2$$

여기에서

$$(\mathrm{Y} - X\beta)/\sigma \sim \mathrm{N_n}(0, \mathrm{I}),\ (\mathrm{I}-\Pi)^2 = (\mathrm{I}-\Pi),\ trace(\mathrm{I}-\Pi) = n-p-1$$ [7)]

이므로 정리 4.4.5의 (b)로부터

$$(n-p-1)\widehat{\sigma^2}/\sigma^2 = (\mathrm{Y} - X\beta)^t (\mathrm{I}-\Pi)(\mathrm{Y} - X\beta)/\sigma^2 \sim \chi^2(n-p-1)$$

예 4.4.9 단순선형회귀모형에서의 표본분포:

예 4.4.7, 예4.4.8과 정리 4.4.6으로부터 단순선형회귀모형에서는

$$\hat{\beta} = \begin{pmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{pmatrix} \sim \mathrm{N}\left(\begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}, \frac{\sigma^2}{\mathrm{S}_{xx}} \begin{pmatrix} \sum_{i=1}^n x_{i1}^2/n & -\overline{x_1} \\ -\overline{x_1} & 1 \end{pmatrix}\right)$$

$$(n-2)\widehat{\sigma^2}/\sigma^2 = \sum_{i=1}^n (\mathrm{Y}_i - \overline{\mathrm{Y}} - \hat{\beta}_1(x_{i1} - \overline{x_1}))^2/\sigma^2 \sim \chi^2(n-2)$$

이고 $\hat{\beta}$과 $\widehat{\sigma^2}$이 독립이다. 따라서 t 분포의 대의적 정의로부터

$$\frac{\hat{\beta}_1 - \beta_1}{\sqrt{\widehat{\sigma^2}/\mathrm{S}_{xx}}} = \frac{(\hat{\beta}_1 - \beta_1)/\sqrt{\sigma^2/\mathrm{S}_{xx}}}{\sqrt{(n-2)(\widehat{\sigma^2}/\sigma^2)/(n-2)}} \sim t(n-2)$$

이 경우에 평균오차제곱합은 다음의 공식을 이용하여 계산할 수 있다.

$$\widehat{\sigma^2} = \sum_{i=1}^n (\mathrm{Y}_i - \overline{\mathrm{Y}} - \hat{\beta}_1(x_{i1} - \overline{x_1}))^2/(n-2) = \left\{\mathrm{S_{YY}} - (\mathrm{S}_{x\mathrm{Y}})^2/\mathrm{S}_{xx}\right\}/(n-2)$$

$$\mathrm{S_{YY}} = \sum_{i=1}^n (\mathrm{Y}_i - \overline{\mathrm{Y}})^2$$

---

7) 행렬의 덧셈과 곱셈의 정의로부터 대각합에 대한 다음 성질이 성립하는 것은 명백하다.

$$trace(A+B) = trace(A) + trace(B),\ trace(AB) = trace(BA)$$

$$\therefore\ trace(\mathrm{I}_{n\times n} - X(X^t X)^{-1} X) = n - trace((X^t X)^{-1} X^t X) = n-p-1$$