다변량 정규분포

정의

- *표준 다변량 정규분포*

1. $z_1, \cdots z_n$을 iid이고 N(0,1)을 따르는 확률변수라고 할 때

1) Z의 결합밀도함수는 iid에서의 조건에 따라

(1) $f_z\,(z) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}}\ \exp\left(-\frac{z^2}{2}\right)\ =\ (\frac{1}{2\pi})^{\frac{n}{2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{n} {z_i}^2\right)$

(2) 위 식을 벡터 형식으로 표현하면

 - $(\frac{1}{2\pi})^{\frac{n}{2}} \exp\left(-\frac{1}{2} z^T z\right)$

정의

- 표준 다변량 정규분포

2) 이 때, Z의 평균과 공분산 행렬은

(1) E(z) = $\prod_{i=1}^{n} E(x_i)$=0

(2) var(z) = $\sum_{i=1}^{n} var(z_i) + 0$ = $\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix}$ = I_z

3) 이 때, Z의 mgf를 구하면

(1) $M_z(t) = E[\exp(t^2 z)]$= $E[\prod_{i=1}^{n} \exp(t_i z_i)]$= $\prod_{i=1}^{n} E[\exp(t_i z_i)]$

(2) 한편, $E[\exp(tz)] = \exp(\frac{1}{2}t^2)$ 이므로

- $\prod_{i=1}^{n}[\exp(\frac{1}{2}t_i{}^2)]$ = $\exp(\frac{1}{2}\sum_{i=1}^{n} t_i{}^2)$ = $\exp(\frac{1}{2}t^T t)$

정의

- 다변량 정규분포

1. Σ를 양반정치이고 nxn의 행렬이라고 한다면, $\Sigma = \Gamma^T \wedge \Gamma$로 분해할 수 있다.

1) 이 때,
(1) ∧는 고윳값을 대각성분으로 가지는 고윳값행렬이고,

(2)Γ는 고유벡터인 정규직교벡터로 이루어진 행렬이다.

2) 이 때, 정규직교행렬의 성질에 따라

(1) $\Gamma^T \Gamma = \mathrm{I}$ 이므로, Σ= $\Gamma^T \wedge^{\frac{1}{2}} \Gamma \ \Gamma^T \wedge^{\frac{1}{2}} \Gamma$로 다시 표현할 수 있다.

(2) 즉, $\Sigma^{\frac{1}{2}}$=$\Gamma^T \wedge^{\frac{1}{2}} \Gamma$ 이다.

정의

- 다변량 정규분포

3) 이제, 행렬 X를 다음과 같이 나타내면

(1) X = $\Sigma^{\frac{1}{2}} z + \mu$, 이 때의 기댓값은 다음과 같다.

(2) E(X) = E[$\Sigma^{\frac{1}{2}} z + \mu$]=$\Sigma^{\frac{1}{2}} \cancel{E[z]} + \mu = \mu$

(3) Var(x) = Var($\Sigma^{\frac{1}{2}} z + \mu$) = var($\Sigma^{\frac{1}{2}} z$) = Σvar(z) = Σ

4) 이 때의 mgf는

(1) E($e^{t^T x}$) = E[$e^{t^T(\Sigma^{\frac{1}{2}} z + \mu)}$]=E[$e^{t^T \Sigma^{\frac{1}{2}} z + t^T \mu}$]= $e^{t^T \mu}$ ·E[$e^{t^T \Sigma^{\frac{1}{2}} z}$]

(2) $t^T \Sigma^{\frac{1}{2}}$ = w로 치환하면, E[$e^{w^T z}$]= $\frac{1}{2} e^{w^T w}$= $\frac{1}{2} e^{(t^T \Sigma^{\frac{1}{2}})^T (t^T \Sigma^{\frac{1}{2}})}$

(3) $e^{t^T \mu} \frac{1}{2} e^{t^T \Sigma t}$

정의

2. 다변량 정규분포의 pdf는 다음과 같다.

1) Z = $(\mathrm{x}-\mu)\cdot\Sigma^{-\frac{1}{2}}$ 이므로

2) |J| = $\Sigma^{-\frac{1}{2}}$, 따라서

$$f(x\,)=\begin{cases}(\frac{1}{2\pi|\Sigma|})^{\frac{n}{2}}\exp\left(-\frac{1}{2}(\mathrm{x}-\mu)^{T}\cdot\Sigma^{-1}(\mathrm{x}-\mu)\right)\\ 0, \qquad else\end{cases}$$

3) 이 때, z는 N(Aμ+b,$\mathrm{A}\Sigma A^{T}$)를 따른다.

정의

3. 다변량 정규분포는 다음과 같이 분할될 수 있다.

1) 벡터 X를 m차원의 벡터라고 할 때,

2) X_1을 n차원 부분벡터로, X_2를 그 나머지 부분벡터로 한다면

(1)

$$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ b_1 \\ b_2 \end{bmatrix} = \begin{matrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} \cdot [\mu_{x_1}] \cdot [\Sigma_{x_1} \quad \Sigma_{x_1 x_2}] \\ \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \cdot [\mu_{x_2}] \cdot [\Sigma_{x_2 x_1} \quad \Sigma_{x_2}] \end{matrix}$$

정의

- 다변량 정규분포

3) 분할은 다음과 같은 방법으로 한다.

(1) X_1=n차원이면 , X_2=m-n=p 차원으로 둔다.

(2) 분할행렬 A를 다음과 같이 정의한다.

$$- \ \mathrm{A} = [I_m \ \vdots \ O_{mp}] = \left[\begin{array}{cccc|cc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{array}\right] \cdot \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ 0 \\ 0 \end{bmatrix} = X_1$$

- 이 때, 분할된 X_1은 평균 μ_{x_1}, 분산 Σ_{x_1}을 따르는 N(μ_{x_1}, Σ_{x_1})을 따른다.

정의

- 다변량 정규분포의 기타성질

2. 정규분포의 결합분포에서 그 조건부 분포는 N($\mu_2 - \Sigma_{12}\Sigma_{22}{}^{-1}(x_2 - \mu_2), \Sigma_{11} - \Sigma_{12}\Sigma_{22}{}^{-1}\ \Sigma_{21}$)을 따른다.

1) 우선, $\mathrm{w} = x_1 - \Sigma_{12}\Sigma_{22}{}^{-1}x_2$ 와 x_2를 가정하면

(1) $\begin{bmatrix} w \\ x_1 \end{bmatrix} = \begin{bmatrix} I & -\Sigma_{12}\Sigma_{22}{}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ 이다.

2) 한편, Y = $Ax + b$일 때, 그 결합분포는 N(Aμ+b,$A\Sigma A^T$)를 따른다.

3) 따라서, 이 결합분포는

(1) $\mu = \begin{bmatrix} I & -\Sigma_{12}\Sigma_{22}{}^{-1} \\ 0 & I \end{bmatrix} \cdot \vec{\mu} = \begin{bmatrix} \mu_1 - \Sigma_{12}\,\Sigma_{22}{}^{-1}\mu_2 \\ \mu_2 \end{bmatrix}$

(2) $\sigma = \begin{bmatrix} I & -\Sigma_{12}\Sigma_{22}{}^{-1} \\ 0 & I \end{bmatrix} \cdot \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \cdot \begin{bmatrix} I & -\Sigma_{12}\Sigma_{22}{}^{-1} \\ 0 & I \end{bmatrix} = \begin{bmatrix} \Sigma_{11} - \Sigma_{12}\,\Sigma_{22}{}^{-1}\Sigma_{21} & 0 \\ 0 & \Sigma_{22} \end{bmatrix}$

(3) 이 때, 대각성분이 0이므로 임의의 벡터 w와 x_2는 독립이다.

예제

- X,Y가 $\mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}$, Σ= $\begin{bmatrix} {\sigma_x}^2 & \sigma_{xy} \\ \sigma_{xy} & {\sigma_y}^2 \end{bmatrix}$ 을 따를때

1) Σ를 다시 쓰면 $\begin{bmatrix} {\sigma_x}^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_x\sigma_y & {\sigma_y}^2 \end{bmatrix}$

2) $\boldsymbol{\Sigma}$의 행렬식 $|\Sigma|$는

(1) ${\sigma_x}^2{\sigma_y}^2$-$(\rho\sigma_x\sigma_y)^2$ = $(\sigma_x\sigma_y)^2$(1- ρ) 이다.

(2) Σ의 역행렬 Σ^{-1}은 $\frac{1}{(\sigma_x\sigma_y)^2(1-\ \rho)}\begin{bmatrix} {\sigma_y}^2 & -\rho\sigma_x\sigma_y \\ -\rho\sigma_x\sigma_y & {\sigma_x}^2 \end{bmatrix}$

예제

- X,Y가 $\mu = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}$, Σ= $\begin{bmatrix} {\sigma_x}^2 & \sigma_{xy} \\ \sigma_{xy} & {\sigma_y}^2 \end{bmatrix}$ 을 따를때

3) 이를 앞서 구한 다변량 정규분포 pdf에 쓰면

(1) $\frac{1}{(2\pi)^{\frac{n}{2}}\,\sigma_x\sigma_y\sqrt{(1-\rho^2)}}\ \exp\left(-\frac{1}{2}(\mathrm{x}-\mu)^{T}\cdot\Sigma^{-1}(\mathrm{x}-\mu)\right)$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}\,\sigma_x\sigma_y\sqrt{(1-\rho^2)}}\ \exp\left(-\frac{1}{2}q\right)$$

(2) 이때 $\mathrm{q} = \frac{1}{1-\rho^2}[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{x-\mu_y}{\sigma_y}\right)^2]$

- 만약 x와 y가 서로 독립일 경우 ρ=0 이고, 따라서
- $\frac{1}{(2\pi)^{\frac{n}{2}}\,\sigma_x\sigma_y}\ \exp\left(-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + \left(\frac{x-\mu_y}{\sigma_y}\right)^2\right)$

예제

- 어떤 부부의 모집단에서 남편의 키를 x_1, 아내의 키를 x_2로 할 때 $\mu_1 = 5.8, \mu_2 = 5.3, \sigma_1 = 0.2, \sigma_2 = 0.2, \rho$=0.6인 이변량 정규분포를 따른다고 한다.

1. x_1 = 6.3일 때 x_2의 조건부 분포를 구하시오

1) 이변량 조건부 분포는 N($\mu_2 - \Sigma_{12}\Sigma_{22}{}^{-1}(x_1 - \mu_1), \Sigma_{11} - \Sigma_{12}\Sigma_{22}{}^{-1}\ \Sigma_{21}$)를 따르므로

(1) μ=5.3 – $\frac{0.6(0.2\cdot 0.2)}{0.2\cdot 0.2}(6.3\ - 5.8)$=5.6

(2) σ= $\sqrt{0.2^2 \text{-}\ \frac{(0.6\cdot 0.2\cdot 0.2)^2}{0.2\cdot 0.2}}$= $0.2\sqrt{1-(0.6)^2}$ = 0.16

2. 위 분포에서 아내의 키가 5.28에서 5.92피트 사이에 위치할 확률을 구하시오

1) $P[\frac{(5.28-5.6)}{0.16} < x_2 < \frac{(5.92-5.6)}{0.16}]$ = $P[-2 < x_2 < 2]$