1. AULA 8

Normal multivariada.

No que segue consideramos a extensão da definição da distribuição normal para o caso multivariado. Consideremos as variáveis aleatórias, $Z_1, Z_2, ..., Z_n$, independentes e identicamente distribuidas com distribuição normal padrão e os números reais $a_{ij}, \mu_j, \quad 1 \leq i \leq n, \quad 1 \leq j \leq n$.

As combinações lineares na forma

$$X_j = a_{1j}Z_1 + a_{2j}Z_2 + ... + a_{nj}Z_n + \mu_j, \quad 1 \leq j \leq n,$$

tem distribuição normal

$$X_j \sim N(\mu_j, \Sigma_{k=1}^n a_{kj}^2).$$

A covariância entre X_i e X_j é

$$Cov(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)] =$$

$$E[(a_{1i}Z_1 + a_{2i}Z_2 + ... + a_{nj}Z_n)(a_{1j}Z_1 + a_{2j}Z_2 + ... + a_{nj}Z_n)] =$$

$$\Sigma_{k=1}^n \Sigma_{l=1}^n a_{ki}a_{lj}E[Z_kZ_l] = \Sigma_{k=1}^n a_{ki}a_{kj}.$$

Utilizando uma notação em forma matricial, fazemos $\mu = (\mu_1, \mu_2, ..., \mu_n)$, $\mathbf{Z} = (Z_1, Z_2, ..., Z_n)$, e a matriz de ordem nxn, $A = (a_{ij})$, temos o sistema de equações

$$\mathbf{X} = \mathbf{Z}A + \mu.$$

A matriz de ordem nxn das covariâncias, $(Cov(X_i, X_j))$, é denotada por $\sum = A'A$, onde $(A'A)_{ij} = \Sigma_{k=1}^n a_{ki}a_{kj}$ e A' é a matriz transposta de A.

Com tais notações definimos

Definição 1.1. Sejam $Z_1, Z_2, ..., Z_n$, independentes e identicamente distribuidas com distribuição normal padrão e seja $\mathbf{X}$ o vetor aleatório obtido de $\mathbf{Z} = (Z_1, Z_2, ..., Z_n)$ através da transformação

$$\mathbf{X} = \mathbf{Z}A + \mu,$$

onde A é uma matriz de números reais de ordem nxn e μ é um vetor real n-dimensional. Então dizemos que $\mathbf{X}$ tem distribuição normal n-variada, com média μ e matriz de covariâncias Σ. $\mathbf{X} \sim N(\mu, \sum)$

Observação 1.2. Desde que $Cov(X_i, X_j) = Cov(X_j, X_i)$, a matriz de covariâncias é simétrica. Se $\mathbf{x}$ é um vetor, n-dimensional de números reais temos

$$\mathbf{x} \sum \mathbf{x}' = \mathbf{x}A'A\mathbf{x}' = (\mathbf{x}A')(\mathbf{x}A')' = |(\mathbf{x}A')|^2 \geq 0$$

e a matriz de covariâncias é definida não negativa.

Se uma matriz de ordem nxn, $\sum$ é definida não negativa, então possui a representação $\sum = A'A$ para alguma matriz de ordem nxn A. Portanto concluimos que toda matriz definida não negativa é matriz de covariâncias de um vetor normal multivariado, pois basta utilizar a matriz A na definição acima.

Proposição 1.3. *Uma matriz, de ordem nxn é matriz de covariâcias de algum vetor normal n-variado se, e somente se, é definida não negativa.*

Observação 1.4. Quando a matriz A tem inversa A^{-1}, a transformação $\mathbf{X} = \mathbf{Z}A + \mu$ é uma correspondência biunívoca de R^n em R^n, com inversa $\mathbf{z} = (\mathbf{x} - \mu)A^{-1}$ e jacobiano $J = Det(A)^{-1}$. Temos

$$\Sigma_{l=1}^{n} z_i^2 = \mathbf{z}\mathbf{z}' = (\mathbf{x} - \mu)A^{-1}(A^{-1})'(\mathbf{x} - \mu)' =$$

$$(\mathbf{x} - \mu)A^{-1}(A')^{-1}(\mathbf{x} - \mu)' = (\mathbf{x} - \mu)(A'A)^{-1}(\mathbf{x} - \mu)'.$$

Portanto $\mathbf{X}$ tem função densidade de probabilidade

$$f_{\mathbf{X}}(x_1, x_2, ..., x_n) = (\frac{1}{\sqrt{2\pi}})^n e^{-\frac{(\mathbf{x}-\mu)(A'A)^{-1}(\mathbf{x}-\mu)'}{2}} |Det(A)^{-1}| =$$

$$(\frac{1}{\sqrt{2\pi}})^n e^{-\frac{(\mathbf{x}-\mu)(\sum)^{-1}(\mathbf{x}-\mu)'}{2}} \sqrt{Det(\sum)^{-1}},$$

para $\mathbf{x} = (x_1, x_2, ..., x_n) \in R^n$, $\sum = AA'$ e $Det(\sum) = Det(A'A) = Det(A)^2 > 0$.

Observação 1.5. Se A é ortonormal, $A'A = I$, a matriz identidade, então $\sum = I$, $Det(\sum) = 1$ e a densidade conjunta de $\mathbf{X}$ é

$$f_{\mathbf{X}}(x_1, x_2, ..., x_n) = (\frac{1}{\sqrt{2\pi}})^n e^{-\frac{(\mathbf{x}-\mu)(\mathbf{x}-\mu)'}{2}} =$$

$$(\frac{1}{\sqrt{2\pi}})^n e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \mu_i)^2}, \quad (x_1, x_2, ..., x_n) \in R^n.$$

Neste caso a densidade é o produto e concluímos que $X_1, X_2, ..., X_n$ são independentes e identicamente distribuidas com distribuição $N(\mu_j, 1)$.

Exemplo 1.6. Sejam $\mu = (0, 0)$ e a matriz A de ordem $2x2$

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Como $AA' = I$, A é ortonormal. Então se Z_1 e Z_2 são independentes com distribuição normal padrão , temos que $X_1 = \frac{Z_1}{\sqrt{2}} + \frac{Z_2}{\sqrt{2}}$ e $X_2 = \frac{Z_1}{\sqrt{2}} - \frac{Z_2}{\sqrt{2}}$ são independentes com distribuição normal padrão. Podemos concluir que $Z_1 + Z_2$ e $Z_1 - Z_2$ são independentes e identicamente distribuidas com distribuição normal com média 0 e variância 2.

Observação 1.7. Se A é ortogonal, isto é, $A'A$ é uma matriz diagonal com elementos diagonais $d_1, d_2, ..., d_n, \quad d_i > 0, 1 \leq i \leq n$, então $\sum = \pi_{i=1}^n d_i$ e a densidade conjunta é

$$f_{\mathbf{X}}(x_1, x_2, ..., x_n) = (\frac{1}{\sqrt{2\pi}})^n \frac{1}{\pi_{i=1}^n d_i} e^{-\frac{1}{2}\sum_{i=1}^n \frac{(x_i-\mu_i)^2}{d_i}}$$

onde $\sum^{-1}$ também é matriz diagonal, tendo elementos $\frac{1}{d_1}, \frac{1}{d_2}, ..., \frac{1}{d_n}$.
Neste caso a densidade é o produto e concluímos que $X_1, X_2, ..., X_n$ são independentes e identicamente distribuidas com distribuição $N(\mu_j, d_j)$.

Observação 1.8. Se a matriz A não é inversível, o jacobiano é nulo e não existe uma função densidade de probabilidade. Dizemos que $\mathbf{X}$ tem temm distribuição normal n-variada degenerada.

A função geradora de momentos conjunta de $\mathbf{X}$ é definida por

$$M_{\mathbf{X}}(t_1, t_2, ..., t_n) = E[e^{(t_1X_1 + t_2X_2 + ... + t_nX_n)}.$$

Como $\sum_{i=1}^n t_iX_i$ é uma combinação linear das variáveis aleatórias independentes e identicamente distribuidas com distribuição normal padrão, também é normalmente distribuida com média $E[\sum_{i=1}^n t_iX_i] = \sum_{i=1}^n t_i\mu_i$ e variância

$$Var(\sum_{i=1}^n t_iX_i) = Cov(\sum_{i=1}^n t_iX_i, \sum_{j=1}^n t_jX_j) = \sum_{i=1}^n \sum_{j=1}^n t_it_jcov(X_i, X_j).$$

Como a função geradora de momentos de uma variável aleatória Y com distribuição $N(\mu, e^2)$ no ponto 1 é $M_Y(1) = e^{\mu + \frac{e^2}{2}}$, temos

$$M_{\mathbf{X}}(t_1, t_2, ..., t_n) = e^{\{\sum_{i=1}^n t_i\mu_i + \frac{1}{2}\sum_{i=1}^n \sum_{j=1}^n t_it_jcov(X_i, X_j)\}}.$$

o que mostra que a distribuição conjunta de $\mathbf{X}$ é completamente determinada a partir das médias $\mu_i, \quad 1 \leq i \leq n$ e das covariâncias $Cov(X_i, X_j) \quad 1 \leq i \leq n$.