BAB III

FUNGSI DISTRIBUSI NORMAL

3.1. DISTRIBUSI NORMAL UNIVARIAT

DEFINISI 3.1.1.

Fungsi densitas untuk variabel acak X normal adalah :

$$f(X) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp^{\{-1/2(X-\mu)^2/\sigma^2\}} \quad , -\infty < X < \infty$$

dengan mean $\mu = E\ (X)$ dan variansi $\sigma^2 = Var\ (X)$, dan dinyatakan sebagai :

$$X \sim N\ (\mu, \sigma^2)$$

Untuk mengestimasi parameter dari fungsi densitas dalam defenisi 3.1.1., yaitu untuk mengestimasi μ dan σ^2, misalkan kita mempunyai sampel acak dari n pengamatan X_1, X_2, ... , X_n diperoleh dari populasi Normal. Maka estimasi dari mean dan variansi adalah :

$$\hat{\mu} = \bar{X}$$

$$= \frac{1}{n} \sum_{i=1}^{n} X_i$$

$$\hat{\sigma}^2 = S^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Apabila mean dari sampel mempunyai nilai yang sama dengan parameter populasi atau mean populasi, maka statistik disini disebut sebagai estimasi tak bias dari parameter, kalau tidak sama maka disebut sebagai estimasi bias dari parameter.

DEFINISI 3.1.2.

Misalkan dari suatu populasi ditarik sampel acak berukuran n dengan variabel acak yang saling bebas $\{X_1, X_2, \ldots, X_n\}$. Diketahui bahwa frekwensi timbulnya X_i ialah $f(X_i, \theta_1, \theta_2, \ldots, \theta_k)$ disingkat dengan $f(X_i/\theta)$. Fungsi ini merupakan fungsi densitas bersama yang dapat ditulis :

$$L(X_1, X_2, \ldots, X_n/\theta) = f(X_1/\theta).f(X_2/\theta) \ldots f(X_n/\theta)$$

$$= \prod_{i=1}^{n} f(X_i/\theta)$$

dimana θ = parameter yang tidak diketahui

Jika variabel acak berdistribusi Normal dengan mean μ dan variansi σ^2, maka dari definisi 3.1.2. kita dapatkan :

$$L = L(X_1, X_2, \ldots, X_n/\mu, \sigma^2)$$

$$= \prod_{i=1}^{n} f(X_i/\mu, \sigma^2)$$

$$= \prod_{i=1}^{n} \frac{1}{(2\pi\sigma^2)^{1/2}} \exp^{\{-1/2(X_i - \mu)^2/\sigma^2\}}$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp \left\{ -\sum_{i=1}^{n} (X_i - \mu)^2/2\sigma^2 \right\}$$

$$L^* = \ln L$$

$$= \frac{-n}{2} \ln 2\pi - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (X_i - \mu)^2$$

Selanjutnya L^* kita defferensialkan parsial terhadap μ dan σ^2, persamaannya menjadi :

$$\frac{\partial L^*}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \mu) = 0$$

$$\frac{\partial L^*}{\partial \sigma^2} = \frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (X_i - \mu)^2 = 0$$

sehingga didapatkan

$$\frac{1}{\sigma^2} \sum_{i=1}^{n} (X_i - \mu) = 0 \text{ , apabila } \sigma^2 \neq 0$$

$$\sum_{i=1}^{n} (X_i - \mu) = 0$$

$$\sum_{i=1}^{n} X_i - n\mu = 0 \implies \hat{\mu} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

dan $$\frac{-n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (X_i - \mu)^2 = 0$$

$$\frac{1}{2\sigma^4} \sum_{i=1}^{n} (X_i - \mu)^2 = \frac{n}{2\sigma^2}$$

$$n\,\sigma^2 = \sum_{i=1}^{n} (X_i - \mu)^2 \implies \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \mu)^2$$

Sehingga kita dapatkan bahwa :

$$E(\hat{\mu}) = \mu$$

$$E(\hat{\sigma}^2) = \left[\frac{n-1}{n} \right] \sigma^2$$

Dari sini terlihat bahwa $\hat{\mu} = \bar{X}$ adalah estimasi tak bias untuk μ, tetapi $\hat{\sigma}^2$ adalah estimasi bias untuk σ^2.

Untuk estimasi tak bias dari σ^2 kita dapatkan dengan :

$$s^2 = \left[\frac{n}{n-1} \right] \hat{\sigma}^2$$

$$= \frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}$$

3.2. DISTRIBUSI NORMAL MULTIVARIAT

DEFINISI 3.2.1.

Suatu vektor acak $\underset{\sim}{x} = (X_1, X_2, \dots, X_p)'$ dikatakan berdistribusi Normal p variat dengan mean $\underset{\sim}{\mu} = (\mu_1, \mu_2, \dots, \mu_p)'$ dan matriks variansi kovariansi Σ, jika fungsi densitas probabilitasnya dapat dituliskan sebagai :

$$f_{\underset{\sim}{x}}(\underset{\sim}{x}) = \frac{1}{(2\pi)^{p/2} \; (\det \Sigma)^{1/2}} \; \exp^{-1/2 \, K}$$

dimana :

$$K = (\underset{\sim}{x} - \underset{\sim}{\mu})' \, \Sigma^{-1} \, (\underset{\sim}{x} - \underset{\sim}{\mu})$$

Fungsi densitas probabilitas ini dinotasikan dengan Np $(\underset{\sim}{\mu}, \Sigma)$ atau $\underset{\sim}{x} \sim$ Np $(\underset{\sim}{\mu}, \Sigma)$.

Misalkan kita mempunyai matriks data

$$X_{p \times n} = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1n} \\ X_{21} & X_{22} & \dots & X_{2n} \\ \vdots & \vdots & & \vdots \\ X_{p1} & X_{p2} & \dots & X_{pn} \end{bmatrix}$$

dimana $X_{11}, X_{12}, \dots, X_{pn}$ adalah variabel-variabel acak. Misalkan, setiap baris dalam sampel acak matriks X berdistribusi p variat :

$$\begin{aligned} \underset{\sim}{x}_1 &= (X_{11} \quad X_{12} \quad \dots \quad X_{1n}) \\ \underset{\sim}{x}_2 &= (X_{21} \quad X_{22} \quad \dots \quad X_{2n}) \\ &\vdots \\ \underset{\sim}{x}_p &= (X_{p1} \quad X_{p2} \quad \dots \quad X_{pn}) \end{aligned}$$

Data sampel acak matriks X dapat dituliskan dalam bentuk vektor yang dinamakan data sampel acak vektor.

$$\underset{\sim}{x} = \begin{bmatrix} \underset{\sim}{x}_1 \\ \underset{\sim}{x}_2 \\ \vdots \\ \underset{\sim}{x}_p \end{bmatrix} \quad \text{atau} \quad \underset{\sim}{x}' = (\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_p)$$

Mean vektor dari sampel acak matriks X adalah :

$$\bar{\underset{\sim}{x}}' = (\bar{X}_1, \bar{X}_2, \ldots, \bar{X}_p)$$

dimana :

$$\bar{X}_i = \frac{1}{n} \sum_{j=1}^{n} X_{ij} \quad , \; i = 1, 2, \ldots, p$$

Matriks variansi-kovariansi dari sampel acak matriks X adalah :

$$S = \begin{bmatrix} S_{11} & S_{12} & \ldots & S_{1n} \\ S_{21} & S_{22} & \ldots & S_{2n} \\ \vdots & \vdots & & \vdots \\ S_{p1} & S_{p2} & \ldots & S_{pn} \end{bmatrix}$$

dimana variansi sampel untuk variabel ke-i adalah :

$$S_i^2 = S_{ii}$$

$$= \frac{1}{n-1} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)^2 \quad , \; i = 1, 2, \ldots, p$$

Akar variansi sampel, $\sqrt{S_{ii}}$ adalah standard deviasi sampel.

Sedangkan kovariansi sampel untuk variabel acak ke-i dan k adalah :

$$S_{ik} = \frac{1}{n-1} \sum_{j=1}^{n} (X_{ij} - \bar{X}_i)(X_{kj} - \bar{X}_k)$$

$$i = 1, 2, \ldots, p$$

$$k = 1, 2, \ldots, p$$

Karena $S_{ik} = S_{ki}$ untuk semua i dan k, sehingga matriks variansi-kovariansi dari sampel acak matriks X berbentuk simetris dan dapat ditulis :

$$S = \begin{bmatrix} S_1^2 & & & \\ S_{12} & S_2^2 & & \\ \vdots & & & \\ S_{1p} & S_{2p} & \cdots & S_p^2 \end{bmatrix}$$

Mean dan kovariansi sampel acak matriks X diatas digunakan untuk mengestimasi mean dan kovariansi populasi dari matriks X.