# BAB II

# TEORI PENUNJANG

## 2.1. Fungsi Pembangkit

1. Fungsi pembangkit satu peubah

Definisi :

Misal X adalah peubah acak dengan distribusi peluang

$P\{ X = k \} = p_k$ ; $k = 0,1,2,\ldots.$

dan $\sum_k p_k = 1$

Jika deret pangkat $G(s) = p_0 + p_1 s + p_2 s^2 + \ldots.$ (2.1.1)

konvergen dalam interval $-s_0 < s < s_0$, maka $G(s)$ dinamakan fungsi pembangkit dari peubah acak X.

Fungsi pembangkit bersifat seragam untuk $|s| < 1$ dan untuk setiap distribusi peluang $\{ p_k \}$ ada satu dan hanya satu fungsi kontinu $G(s)$ dan demikian pula sebaliknya.

Nilai-nilai dari distribusi peluang $\{ p_k \}$ dapat diperoleh dengan mendeferensialkan suku demi suku $G(s)$ hasilnya :

$$p_k = \frac{1}{k!} \left. \frac{d^k G(s)}{ds^k} \right|_{s=0} \quad ; k=0,1,2,\ldots. \qquad (2.1.2)$$

Jelas bahwa distribusi peluang $\{ p_k \}$ merupakan koefisien dari $s^k$ dalam fungsi pembangkit $G(s)$.

Berarti pula deret pangkat (2.1.1) juga mendefinisikan ekspektasi $G(s) = E(s^X)$.

Dari persamaan (2.1.3) kita dapat menggunakan fungsi pembangkit untuk mendapatkan ekspektasi dan momen-momen derajat tinggi dari peubah acak dengan cara diferensiasi.

Misal :

$$E(X) = \frac{d\,G(s)}{ds}\Bigg|_{s=1} \qquad (2.1.4)$$

$$E(X(X-1)) = \frac{d^2G(s)}{ds^2}\Bigg|_{s=1} \qquad (2.1.5)$$

$$E(X(X-1)\ldots(X-n+1)) = \frac{d^nG(s)}{ds^n}\Bigg|_{s=1} \qquad (2.1.6)$$

Variansi diperoleh dari (2.1.4) dan (2.1.5):

$$\sigma_X^2 = E(X(X-1)) + E(X) - (E(X))^2 \qquad (2.1.7)$$

2. Fungsi pembangkit dua peubah

Misalkan $X_1$, $X_2$ dua peubah acak dengan distribusi peluang :

$$P\{X_1 = k_1 , X_2 = k_2\} = P_{k_1 k_2} \; ; \; k_1, k_2 = 0,1,2,\ldots.$$

dan

$$\sum_{k_1} \sum_{k_2} P_{k_1 k_2} = 1$$

maka fungsi pembangkit dari peubah acak $X_1$ dan $X_2$ didefinisikan :

$$G(s_1,s_2) = \sum_{k_1} \sum_{k_2} P_{k_1 k_2} \, s_1^{k_1} \, s_2^{k_2} \qquad (2.1.8)$$

Fungsi pembangkit $G(s_1,s_2)$ bersifat konvergen pada $|s_1| < 1$ dan $|s_2| < 1$.

Distribusi peluang $\{P_{k_1,k_2}\}$ dapat diperoleh dengan mendeferesialkan suku demi suku persamaan (2.1.8).

$$P_{k_1k_2} = \left. \frac{\partial\, G(s_1,s_2)}{k_1!\; k_2!\; \partial s_1^{k_1}\; \partial s_2^{k_2}} \right|_{s_1=0,\; s_2=0} \qquad (2.1.9)$$

Fungsi pembangkit dari distribusi marginal $P\{X_1 = k_1\}$ adalah :

$$E(s_1^{X_1}) = G(s_1,1) \qquad (2.1.10)$$

Fungsi pembangkit dari distribusi marginal $P\{X_2 = k_2\}$ adalah

$$E(s_2^{X_2}) = G(1,s_2) \qquad (2.1.11)$$

Perubah acak $X_1$ dan $X_2$ bebas jika dan hanya jika :

$$E(s_1^{X_1}\; s_2^{X_2}) = G(s_1,s_2) = G(s_1,1)\; G(1,s_2) \qquad (2.1.12)$$

Ekspektasi dan momen-momen derajat tinggi diperoleh dengan cara mendeferensialkan fungsi pembangkit $G(s_1,s_2)$.

$$E(X_i) = \left. \frac{\partial\, G(s_1,s_2)}{\partial s_i} \right|_{s_1=1\;,\; s_2=1} ;\; i=1,2 \qquad (2.1.13)$$

$$E(X_i(X_i-1)) = \left. \frac{\partial^2\, G(s_1,s_2)}{\partial s_i^{\,2}} \right|_{s_1=1\;,\; s_2=1} ;\; i=1,2$$

$$(2.1.14)$$

Variasi diperoleh dari (2.1.13) dan (2.1.14) :

$$\sigma^2_{X_i} = E(X_i(X_i-1)) + E(X_i) - (E(X_i))^2 \; ; \; i = 1,2 \qquad (2.1.15)$$

Kovariansi antara $X_1$ dan $X_2$ adalah :

$$\sigma^2_{X_1X_2} = \left.\frac{\partial^2 G(s_1,s_2)}{\partial s_1 \, \partial s_2}\right|_{s_1=1\,,\; s_2=2} - \left.\left[\frac{\partial G(s_1,s_2)}{\partial s_1}\right]\left[\frac{\partial G(s_1,s_2)}{\partial s_2}\right]\right|_{s_1=1,\; s_2=1} \qquad (2.1.16)$$

## 2.2. Pengujian Chi Kuadrat

Pengujian chi kuadrat adalah pengujian mutu penjajagan, yaitu kita menguji apakah peubah acak X mempunyai distribusi F(x) yang tertentu atau tidak.

Misalkan $X_1$, $X_2$, $X_3$, ... , $X_n$ bebas, masing-masing dengan distribusi tertentu yang sama.

$H_0$ = distribusi $X_i$ adalah F(X)

$H_1$ = tidak demikian.

Data sampel adalah $x_1$, $x_2$, $x_3$, ..., $x_n$.

Distribusi sampel yang dilukiskan sebagai histogram atau poligon frekuensi merupakan suatu bayangan statistik dari distribusi X, jadi dapat dibandingkan dengan fungsi padat F(x).

Dengan jalan ini kita memperoleh secara kualitas persesuaian atau tidak antara kedua distribusi tersebut.

Untuk dapat mengetahui derajat persamaan tersebut kita memerlukan ukuran kuantitas mengenai besarnya deviasi atau penyimpangan dari distribusi hipotesis terhadap distribusi sampel. Jika ukuran ini melampaui suatu nilai kritis, maka kita menolak $H_0$ yaitu hipotesa bahwa distribusi X adalah F(x).

K. Pearson menemukan pengujian mengenai hal di atas yaitu pengujian yang disebut pengujian Chi Kuadrat.

Kita membagi jangkauan dari X menjadi r (hingga) golongan $G_1, G_2, G_3, \ldots, G_n$ yang saling bertentangan dan misalkan $p_i = P(X \in G_i)$ dan $f_i$ adalah frekuensi data sampel yang termasuk $G_i$, maka

$$\begin{aligned} p_1 + p_2 + \ldots + p_r &= 1 \\ f_1 + f_2 + \ldots + f_r &= n \end{aligned} \qquad (2.2.1)$$

Misalkan $H_0$ benar, maka $f_i$ merupakan frekuensi kejadian bahwa $X \in G_i$ dari n pengamatan, mempunyai distribusi binomial $b(n,p_i)$. Menurut dalil de Moivre-Laplace, jika n besar maka distribusi $f_i$ menjadi normal asimtotis $\mathcal{N}(np_i, np_i(1-p_i))$.

Jika $H_0$ benar dan n besar, kita dapat mengharapkan ada persesuaian antara barisan $f_1, f_2, \ldots, f_r$ dengan barisan $np_1, np_2, \ldots, np_r$.

$f_i$ sering juga ditulis sebagai $o_i$ disebut frekuensi pengamatan, dan $np_i$ sering ditulis sebagai $e_i$ disebut frekuensi harapan.

Sebagai ukuran deviasi frekuensi pengamatan $o_i = f_i$ terhadap frekuensi harapan $e_i = np_i$ diambil.

$$\chi^2 = \sum_{1}^{r} \frac{(o_i - e_i)^2}{e_i} = \sum_{1}^{r} \frac{(f_i - np_i)^2}{np_i} \qquad (2.2.2)$$

Misalkan $Z_i = \dfrac{f_i - np_i}{\sqrt{np_i}}$ , maka $Z_i \approx \dfrac{fp_i - np_i}{\sqrt{np_i(1-p_i)}}$

karena $1 - p_i \approx 1$

Jadi distribusi $Z_i$ untuk n besar berbentuk normal asimtotis.

$$\mathcal{N}(0,1) \text{ sehingga } \chi^2 = \sum_{i=1}^{r} Z_i^{\ 2} \qquad (2.2.3)$$

Karena persamaan (2.2.1), maka $\chi^2$ adalah jumlah r peubah acak yang normal asimtotis $\mathcal{N}(0,1)$ memenuhi satu hubungan linier

$$\sum_{i=1}^{r} Z_i \sqrt{p_i} = 0$$

Untuk dapat menemukan distribusi (2.2.3) kita melihat dalil yang menerangkan bahwa

$$\frac{n\ s^2}{\sigma^2} = \sum_{i=1}^{n} \left( \frac{X_i - X}{\sigma} \right)^2$$

mempunyai distribusi $\chi^2$ dengan (n-1) derajat kebebasan, karena $\dfrac{X_i - \overline{X}}{\sigma}$ memenuhi hubungan linier.

Jadi jika $H_0$ benar, $$\chi^2 = \sum_{i=1}^{r} \frac{(f_i - np_i)^2}{np_i} \qquad (2.2.5)$$

mempunyai distribusi $\chi^2$ dengan (r-1) derajat kebebasan.

Jadi $H_0$ diterima jika $\chi^2$ perhitungan < $\chi^2$

$H_0$ ditolak jika $\chi^2$ perhitungan > $\chi^2$

dimana $P(\chi > \chi_p^{\,2}) = P/100 = \alpha$ ($\alpha$ = taraf kepercayaan).

Dalam pengujian mutu penjajagan Chi Kuadrat $\chi^2$ jika ada parameter yang ditaksir, distribusi asimtotis $\chi^2$ yang telah diterangkan masih berlaku hanya derajat kebebasan diganti dengan (r-s-1) dengan r menyatakan banyaknya golongan dan s menyatakan banyaknya parameter yang ditaksir.

Dalam pengujian kecocokan Chi Kuadrat $\chi^2$ pada tabel kontingensi k x h, distribusi asimtotis $\chi^2$ tetap berlaku, hanya derajat kebebasan diganti dengan (k-1)x(h-1) dengan k menyatakan banyaknya baris dan h banyaknya kolom.