BAB II

DASAR TEORI

2.1. Probabilitas

Peristiwa pasti S adalah peristiwa yang terjadi pada setiap usaha. Gabungan $A + B$ dua peristiwa $A + B$ adalah peristiwa yang terjadi bila A atau B atau kedua-duanya terjadi. Irisan AB peristiwa-peristiwa A da B adalah peristiwa yang terjadi bila kedua peristiwa A dan B terjadi. Peristiwa-peristiwa A dan B disebut saling asing bila terjadinya satu dari mereka meniadakan terjadinya peristiwa yang lain (Athanasios, terjemahan 1985).

Definisi 2.1.1.

Probabilitas peristiwa A yaitu $P(A)$ adalah bilangan tak negatif yang diterapkan pada peristiwa ini.

$$P(A) \geq 0.$$

Probabilitas peristiwa pasti sama dengan 1

$$P(S) = 1.$$

Bila peristiwa A dan B saling asing, maka

$$P(A + B) = P(A) + P(B).$$

Definisi 2.1.2.

Probabilitas peristiwa A ditentukan apriori tanpa pelaksanaan eksperimen yang sebenarnya. Probabilitas diberikan oleh rasio

$$P(A) = \frac{N_A}{N}$$

dimana N jumlah hasil yang mungkin dan N_A jumlah hasil yang sesuai dengan peristiwa A.

Kita dapat menentukan probabilitas tak hanya pada gabungan dan irisan berhingga peristiwa-peristiwa, tetapi juga limitnya.

Definisi 2.1.3.

Bila peristiwa $A_1, A_2, \ldots, A_n$ saling asing maka

$$P(A_1 + A_2 + \ldots + A_n) = P(A_1) + P(A_2) + \ldots + P(A_n),$$

Bila ruang S terdiri dari N hasil dan n bilangan berhingga, maka probabilitas semua peristiwa dapat dinyatakan dalam bentuk probabilitas $P\{\zeta\} = p_i$ peristiwa elementer $\{\zeta_i\}$. Tentu saja dari aksioma terlihat bahwa bilangan p_i harus tak negatif dan jumlahnya harus sama dengan satu :

$$p_i \geq 0, \qquad p_1 + p_2 + \ldots + p_N = 1.$$

Definisi 2.1.4.

Probabilitas bersyarat peristiwa A dengan diketahui M, ditulis $P(A|M)$ didefinisikan sebagai

$$P(A|M) = \frac{P(AM)}{P(M)},$$

dimana diandaikan bahwa $P(M)$ tidak nol.

Sifat-sifat:

1. Bila $M \subset A$ maka $P(A|M) = 1$

   bentuk diatas terbukti karena $AM = M$.

2. Bila $A \subset M$ maka $P(A|M) = \dfrac{P(AM)}{P(M)} \geq P(A)$.

Teorema 2.1.1.

Jika N_A, N_M dan N_{AM} masing-masing adalah jumlah terjadinya peristiwa-peristiwa A, M dan AM maka

$$P(A|M) = \frac{N_{AM}}{N_M}.$$

Bukti.

$P(A) = \frac{N_A}{N}$, $P(M) = \frac{N_M}{N}$, dan $P(AM) = \frac{N_{AM}}{N}$, karena itu

$$P(A|M) = \frac{P(AM)}{P(M)} = \frac{N_{AM}/N}{N_M/N} = \frac{N_{AM}}{N_M}.$$

Hasil ini dapat dinyatakan sebagai berikut : Bila kita menghilangkan seluruh usaha dimana peristiwa M tidak terjadi dan kita hanya memakai barisan bagian usaha dimana M terjadi, maka $P(A|M)$ sama dengan frekuensi relatif terjadinya N_{AM}/N_M peristiwa A pada barisan bagian tersebut.

Teorema 2.1.2.

Jika peristiwa A dan B saling asing, maka

$$P(A+B|M) = P(A|M) + P(B|M).$$

Bukti.

A dab B saling asing, maka peristiwa AM dan BM juga saling asing, karena itu

$$P(A|M) = \frac{P[(A+B)M]}{P(M)}$$

$$= \frac{P(AM)+P(BM)}{P(M)}$$

$$= \frac{P(AM)}{P(M)} + \frac{P(BM)}{P(M)}$$

$$= P(A|M) + P(B|M).$$

Dari hasil diatas terlihata bahwa seluruh hasil mengenai probabilitas juga berlaku untuk probabilitas bersyarat.

Definisi 2.1.5.

Dua peristiwa A dan B disebut independen bila

$$P(AB) = P(A)P(B).$$

Misalkan N_A, N_B dan N_{AB} masing-masing menyatakan jumlah terjadinya peristiwa A, B dan AB, maka

$$P(A) = \frac{N_A}{N},\ P(B) = \frac{N_B}{N}\ ,\text{dan}\ P(AB) = \frac{N_{AB}}{N_B}.$$

Bila peristiwa A dan B independen, maka frekuensi relatif N_A/N dari terjadinya A pada bagian N usaha awal sama dengan frekuensi relatif N_{AB}/N_B dari terjadinya A pada barisan bagian dimana B telah terjadi.

Akan diperlihatkan bahwa bila peristiwa A dan B independen, maka peristiwa $\overline{A}$ dan B serta peristiwa $\overline{A}$ dan $\overline{B}$ juga independen. Seperti telah diketahui peristiwa AB dan $\overline{A}\,B$ saling asing dan

$$B = AB + \overline{A}\,B \qquad P(\overline{A}) = 1 - P(A)$$

dari ini dan definisi 2.1.5 terlihat bahwa

$$\begin{aligned} P(\overline{A}\,B) &= P(B) - P(AB) \\ &= [1-P(A)]P(B) \\ &= P(\overline{A})P(B). \end{aligned}$$

Definisi 2.1.6.

Tiga peristiwa A_1, A_2, A_3 disebut saling independen bila A_1, A_2, A_3 independen dalam pasangan :

$$P(A_i A_j) = P(A_i)P(A_j) \qquad i \neq j$$

dan

$$P(A_1 A_2 A_3) = P(A_1)P(A_2)P(A_3).$$

Disini perlu ditekankan bahwa tiga peristiwa independen dalam pasangan tetapi tidak independen.

Contoh 2.1.

Misalkan probabilitas A, B, dan C mempunyai probabilitas sama $P(A) = P(B) = P(C) = 1/5$ dan irisan AB, AC, BC dan ABC juga mempunyai probabilitas sama

$$p = P(AB) = P(AC) = P(BC) = P(ABC)$$

(a) Bila $p = 1/25$, maka ketiga peristiwa independen dalam pasangan tetapi tidak independen karena

$$P(ABC) \neq P(A)P(B)P(C).$$

(b) Bila $p = 1/125$, maka $P(ABC) = P(A)P(B)P(C)$ tetapi ketiga peristiwa tidak independen karena

$$P(AB) \neq P(A)P(B).$$

Dari independensi peristiwa A, B, dan C terlihat bahwa :

1. Setiap peristiwa independen dengan irisan dua peristiwa lainnya. Dari (definisi 2.1.6) terlihat bahwa

$$P(A_1 A_2 A_3) = P(A_1)P(A_2)P(A_3) = P(A_1)P(A_2 A_3)$$

karena itu, peristiwa A_1 dan A_2A_3 independen.

2. Bila kita mengganti satu atau beberapa peristiwa dengan komplemennya, maka peristiwa hasilnya juga independen. Karena $A_1A_2 = A_1A_2A_3 + A_1A_2\overline{A}_3$ dan $P(\overline{A}_3) = 1 - P(A_3)$ maka bersama (ketentuan 1) dapat disimpulkan bahwa

   $P(A_1A_2\overline{A}_3) = P(A_1A_2) - P(A_1A_2)P(A_3) = P(A_1)P(A_2)P(\overline{A}_3)$.

   Oleh karena itu, peritiwa A_1 dan A_2 dan A_3 memenuhi (definisi 2.1.6) dan seperti telah dibuktikan sebelumnya ketiga peristiwa itu independen dalam pasangan.

3. Setiap peristiwa independen dengan gabungan dua peristiwa yang lain

2.2. Variabel Random dan Nilai Harapan

Variabel random adalah suatu fungsi sederhana dari suatu himpunan hasil untuk bilangan riil. Distribusi untuk suatu variabel random diskrit adalah suatu daftar nilai yang mungkin untuk sepanjang variabel random dengan probabilitas nilai-nilai kejadian.

Definisi 2.2.1.

Sebuah percobaan random dengan ruang sampel ξ. Suatu fungsi X yang mengawankan setiap elemen $C \in \xi$ dengan satu dan hanya satu bilangan riil $X_{(C)} = x$, disebut sebagai variabel random dengan range $A = \{ X : X = X_{(C)}, C \in \xi \}$.

Definisi 2.2.2.

(i). Jika range A dari varibel random X memuat titik-titk yang banyaknya berhingga atau anggota A dapat dipasangkan berkorespondensi satu-satu dengan bilangan integer positif, maka X disebut suatu varibel diskrit.

(ii). Jika range A dari variabel random X berupa interval atau kumpulan dari interval-interval, maka X disebut suatu variabel random kontinu.

Lambang $\sum_i$ menyatakan penjumlahan dalam hal diskrit, sama fungsinya dengan tanda integral dalam hal kontinu. Peubah acak biasanya dinyatakan dalam huruf besar sedangkan nilainya dalam huruf kecil.

Definisi 2.2.3.

(i). Suatu fungsi $f(x)$ yang didefinisikan pada himpunan A ke himpunan bilangan riil disebut sebagai fungsi densitas probabilitas diskrit, jika memenuhi :

i. $f(x) \geq 0,\ \forall\, x \in A.$

ii. $\sum_A f(x) = 1.$

(ii). Suatu fungsi $f(x)$ yang didefinisikan pada himpunan A ke himpunan bilangan riil disebut fungsi densitas probabilitas kontinu, jika memenuhi :

iii. $f(x) > 0,\ \forall\, x \in A.$

iv. $\int_{-\infty}^{+\infty} f(x)\, \mathrm{dx} = 1.$

Definisi 2.2.4.

Misalkan X adalah variabel random yang mempunyai fungsi densitas probabilitas $f(x)$, maka ekspektasi dari variabel random X adalah :

$E(X) = \sum_x x\, f(x)$, untuk variabel random diskrit.

$E(X) = \int_{-\infty}^{+\infty} x\, f(x)\, dx$, untuk variabel random kontinu.

Sifat-sifat ekspektasi

1. $E(a) = a$, bila a adalah konstanta.
2. $E(aX + b) = aE(X) + b$, bila a,b konstanta.
3. $E(\sum k_i X_i) = \sum k_i E(X_i)$.

Definisi 2.2.5.

Diberikan variabel random X dan nilai rata-rata dari X didefinisikan dengan $\mu = E(X)$. Ekspekstasi dari kuadrat antara X dengan μ disebut Var(X) yaitu Var(X) = $E\{[X - \mu]^2\}$.

Sifat-sifat variansi:

1. $\mathrm{Var}(a) = 0$, bila a konstanta.
2. $\mathrm{Var}(aX + b) = a^2\,\mathrm{Var}(X)$, bila a dan b konstanta.

2.3. Distribusi Binomial dan Multinomial

Percobaan binomial adalah suatu percobaan yang terdiri atas beberapa usaha dimana setiap usaha akan memberikan dua hasil yang mungkin, yaitu sukses dan gagal (Ronald E.W, 1995). Suatu percobaan dikatakan sebagai suatu usaha binomial apabila memenuhi syarat-syarat sebagai berikut :

1. Percobaan terdiri atas n usaha yang berulang.
2. Tiap usaha memberikan hasil yang ditentukan dengan sukses atau gagal.
3. Peluang sukses dinyatakan dengan p, peluang gagal dinyatakan dengan $q = 1 - p$.
4. Tiap usaha bebas.

Definisi 2.3.1.

Bila suatu ulangan mempunyai peluang keberhasilan p dan peluang gagal $q = 1 - p$, maka distribusi peluang bagi peubah acak X yaitu banyaknya keberhasilan dalam n ulangan yang bebas adalah :

$$\Pr(X = x) = \binom{n}{x} p^x q^{n-x} \text{ ,untuk } x = 0,1,2,\ldots.,n$$

dimana $\binom{n}{x} = \dfrac{n!}{x(n-x)!}$,

Nilai harapan untuk distribusi binomial adalah:

$$E(X) = \sum_{x=0}^{n} x \binom{n}{x} (1-p)^{n-x}$$

$$= np$$

dan varian untuk distribusi binomial adalah:

$$Var(X) = \sum_{x=0}^{n} (x - np)^2 \binom{n}{x} (1-p)^{n-x}$$

$$= npq .$$

Contoh 2.3.1.

Tentukan peluang mendapatkan tepat tiga bilangan dua bila sebuah dadu seimbang dilemparkan lima kali.

Peluang keberhasilan setiap ulangan yang bebas ini adalah 1/6 dan peluang kegagalan adalah 5/6. Dalam hal ini muncul bilangan dua dianggap keberhasilan,

maka $P(X = 3) = \binom{5}{3} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^2$

$$= \frac{5!}{3!2!} \frac{1}{6^2} \frac{5^2}{6^5}$$

$$= 0{,}032.$$

Seandainya dalam percobaan binomial tersebut setiap ulangan menghasilkan lebih dari kemungkinan dua hasil, maka percobaan itu menjadi apa yang disebut percobaan multinomial (Ronald E.W, 1995). Secara umum, bila ulangan dapat menghasilkan satu diantara k kemungkinan hasil percobaan $E_1, E_2, \ldots, E_k$, masing-masing dengan peluang p_1, $p_2, \ldots$, p_k, maka distribusi multinomial memberikan peluang terjadinya X_1 kali kejadian $E_1, \ldots$ X_k kali kejadian E_k dalam n ulangan yang bebas dengan $X_1 + X_2 + \ldots + X_k = n$. distribusi peluang bersama ini akan dilambangkan dengan $f(X_1, X_2, \ldots X_k\ ;\ p_1, p_2, \ldots, p_k, n)$. Jelas bahwa $p_1 + p_2 + \ldots + p_k = 1$, karena hasil yang muncul dari setiap ulangan pastilah salah satu diantara k kemungkinan hasil.

Definisi 2.3.2.

Bila setiap ulangan menghasilkan salah satu dari k hasil percobaan E_1, $E_2, \ldots$, E_k, dengan peluang p_1, $p_2, \ldots$, p_k distribusi peluang bagi peubah acak X_1, $X_2, \ldots$, X_k yang menyatakan beberapa kali $E_1, E_2, \ldots, E_k$ terjadi dalam n ulangan yang bebas, adalah:

$$\Pr(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k) = \frac{n!}{x_1! x_2 \ldots x_k!} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$$

$$= \frac{n!}{\prod_{i=1}^{k} x_i!} \prod_{i=1}^{k} p_i^{x_k}$$

dengan $\sum_{i=1}^{k} X_i = n$ dan $\sum_{i=1}^{k} p_i = 1$.

Nilai harapan untuk distribusi multinomial adalah:

$$E(Xi) = np_i$$

dan varian untuk distribusi multinomial adalah:

$$Var(X_i) = np_i(1 - p_i).$$

Contoh 2.3.2.

Bila dua dadu dilemparkan 6 kali, berapa peluang mendapatkan jumlah bilangan yang muncul sebesar 7 atau 11 sebanyak dua kali, bilangan yang sama pada kedua dadu sekali, dan kemungkinan lainnya tiga kali.

Jawab.

Dua mata dadu dilemparkan $N = 216$. Jika dimisalkan A kejadian jumlah 7 muncul didapat N_A = 36, B kejadian jumlah 11 muncul didapat N_B = 12, C kejadian jumlah bilangan yang sama muncul didapat Nc = 36, dan D kejadian jumlah kemungkinan lainnya didapat N_D = 132. Sehingga diperoleh:

$$p_1 = P(A) \cup P(B)$$

$$= \frac{36}{216} + \frac{12}{216}$$

$$= \frac{2}{9}$$

$$p_2 = P(C) = \frac{36}{216}$$

$$= \frac{1}{6}$$

$$p_3 = (p_1 + p_2)^c = 1 - p_1 - p_2$$

$$= \frac{11}{18},$$

atau $p_3 = P(D) = \frac{132}{216}$

$$= \frac{11}{18}$$

Dan dimisalkan:

E_1 = jumlah 7 atau 11 muncul

E_2 = jumlah pasangan bilangan yang sama muncul

E_3 = jumlah kemungkinan lainnya selain E_1 dan E_2 diatas.

Dengan menggunakan distribusi multinomial peluang $X_1 = 2$, $X_2 = 1$, dan $X_3 = 3$, adalah:

$$P(X_1 = 2, X_2 = 1, X_3 = 3) = \left(\frac{6!}{2!\,1!\,3!}\right)\left(\frac{2}{9}\right)^2\left(\frac{1}{6}\right)^1\left(\frac{11}{18}\right)^3$$

$$= 60\left(\frac{4}{81}\right)\left(\frac{1}{6}\right)\left(\frac{1331}{5832}\right)$$

$$= 0{,}1127.$$

2.4. Estimasi Maksimum Likelihood

Prosedur taksiran maksimum likelihood menguji apakah taksiran maksimum yang diketahui dari fungsi likelihood suatu sampel nilainya sudah memaksimumkan fungsi itu.

Definisi 2.4.1.

Misalkan $X_1, X_2, \ldots\ldots\ldots, X_n$ adalah sampel random dari populasi dengan distribusi fungsi $f(x,\theta)$, dimana parameter θ tidak diketahui, maka fungsi likelihood sampel tersebut adalah:

$$L(\theta) = f(x_1,\theta).f(x_2,\theta).\ldots\ldots\ldots\ldots f(x_n,\theta).$$

Bila fungsi likelihood terdiferensialkan ke θ, maka taksiran maksimum likelihood yang mungkin adalah:

$$(\hat{\theta}_1, \hat{\theta}_2, \ldots\ldots, \hat{\theta}_k) \ni L(\theta) \text{ maksimum}$$

Untuk membuktikan bahwa $\hat{\theta}_i$ benar-benar memaksimumkan $L(\theta)$ harus ditunjukkan:

$$\left.\frac{\partial L^2(\theta)}{\partial \hat{\theta}_i^{\,2}}\right|_{\theta_i = \hat{\theta}_i} < 0 .$$

Dalam banyak kasus, dimana diferiansi digunakan akan lebih mudah bekerja pada logaritma natural dari $L(\theta)$, yaitu :

$$K(\theta) = \ln L(\theta) .$$

Hal ini dimungkinkan karena fungsi log naik tegas pada (θ, ∞) yang berarti bahwa:

$$L(\theta) \text{ dan } \ln L(\theta)$$

mempunyai ekstrem yang sama. Jelasnya untuk menentukan taksiran maksimum likelihood dari θ, langkah-langkahnya:

1. Tentukan fungsi likelihood $L(\theta) = f(X_1,\theta).f(X_2,\theta)\ldots\ldots f(X_n,\theta)$.
2. Bentuk logaritma natural likelihood $K(\theta) = \ln L(\theta)$.

3. Bentuk persamaan likelihood dan selesaikan.

Contoh 2.4.1.

Misalkan X berdistribusi normal dengan rata-rata μ dan varian σ^2 tidak diketahui. Tentukan taksiran maksimum likelihood untuk μ dan σ^2.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X_i-\mu)^2}{2\sigma^2}}$$

sehingga dapat ditentukan fungsi likelihoodnya yaitu

$$L(\mu,\sigma^2) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(X_i-\mu)^2}{2\sigma^2}}$$

$$= \frac{1}{\sigma\sqrt{2\pi}}\left[e^{-\frac{(X_1-\mu)^2}{2\sigma^2}}\, e^{-\frac{(X_2-\mu)^2}{2\sigma^2}} \,.....\, e^{-\frac{(X_n-\mu)^2}{2\sigma^2}} \right]$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}}\, e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2}.$$

Dengan menggunakan logaritma natural, maka fungsi likelihoodnya menjadi

$$\ln L(\mu,\sigma^2) = \ln\left[\frac{1}{(2\pi\sigma^2)^{n/2}}\, e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2} \right]$$

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

Taksiran maksimum likelihood untuk $\hat{\mu}$ dan $\hat{\sigma}^2$ didapat jika

$$\frac{\partial \ln L(\mu,\sigma^2)}{\partial \hat{\mu}} = 0 \ ; \ \frac{\partial \ln L(\mu,\sigma^2)}{\partial \hat{\sigma}^2} = 0 .$$

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

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

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

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

$$\frac{\partial \ln L(\mu,\sigma^2)}{\partial(\hat{\sigma}^2)^2} = -\frac{n}{2}\left(\frac{1}{\hat{\sigma}^2}\right) + \frac{1}{2\hat{\sigma}^4}\sum_{i=1}^{n}(X_i - \mu)^2 = 0$$

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

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

Untuk membuktikan bahwa $\hat{\mu}$ dan $\hat{\sigma}^2$ benar-benar memaksimumkan L(μ,σ^2)

harus ditunjukkan:

$$\left.\frac{\partial L^2(\mu,\sigma^2)}{(\partial\hat{\mu})^2}\right|_{\mu = \hat{\mu}} < 0 \ ; \ \left.\frac{\partial L^2(\mu,\sigma^2)}{(\partial\hat{\sigma}^2)}\right|_{\sigma^2 = \hat{\sigma}^2} < 0$$

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

$$\frac{\partial L^2(\mu,\sigma^2)}{(\partial\hat{\sigma}^2)^2} = \frac{n}{\hat{\sigma}^3}\sum_{i=1}^{n}(X_i - \hat{\mu})(-1) < 0$$

maka

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

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

Sehingga dihasilkan estimator maksimum likelihood yaitu:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

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