BAB II

MATERI PENUNJANG

2.1. Teori Peluang

Diberikan Ω suatu ruang sampel dan A sebagai suatu koleksi dari setiap himpunan bagian dari Ω.

Definisi 1. Fungsi Probabilitas.

Suatu Fungsi Probabilitas P(•) merupakan suatu fungsi yang ditetapkan dengan domain (daerah asal) A dan co-domain (daerah hasil) [0,1] yang memenuhi aksioma-aksioma sebagai berikut :

(i) $P(A) \geq 0$ untuk setiap $A \in A$,

(ii) $P(\Omega) = 1$

(iii) Jika $A_1, A_2, \ldots$ suatu barisan peristiwa saling asing di dalam A (yaitu $A_i \cap A_j = 0$ untuk $i \neq j$; $j = 1,2, \ldots$)

dan jika $A_1 \cup A_2 \cup A_3 \cup \ldots = \bigcup_{i=1}^{n} A_i \in A$,

maka $P[\bigcup_{i=1}^{n} A_i] = \bigcup_{i=1}^{n} P(A_i)$

Untuk selanjutnya, koleksi dari setiap himpunan bagian dari Ω yaitu A disebut Ruang Peristiwa dari Ω.

Definisi 2. Ruang Probabilitas.

Susunan tripel $\{\Omega, A, P\}$ disebut suatu Ruang Probabilitas, dimana Ω merupakan suatu ruang sampel, A ruang peristiwa dari Ω dan P (•) adalah suatu fungsi probabiitas dengan domain A.

Definisi 3. Partisi.

Suatu Partisi U = [$A_1, A_2, \ldots, A_n$] dari himpunan Ω merupakan koleksi himpunan bagian dari A_i yang memenuhi :

(i) $A_i \subset \Omega$

(ii) $A_i \cap A_j = \emptyset, i \neq j$ dan

(iii) $A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$

Definisi 4. Probabilitas Bersyarat.

Diambil A dan B dua peristiwa dalam A, dari ruang probabilitas yang diberikan $\{\Omega, A, P\}$. Probabilitas bersyarat dari A diberikan B dinotasikan dengan $P(A|B)$, dan didefinisikan :

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{dimana } P(B) > 0$$

Teorema 1. Probabilitas Total.

Diberikan suatu ruang probabilitas $\{\Omega, A, P\}$. Jika $A_1, A_2, \ldots, A_n$ merupakan koleksi himpunan peristiwa saling asing di dalam A yang memenuhi $\Omega = \bigcup_{i=1}^{n} A_i$ dan $P(A_i) > 0$ untuk i = 1, 2, ..., n, maka untuk

setiap $B \in A$ dimana $P(B) > 0$ berlaku

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

Bukti :

Sebagaimana diketahui bahwa untuk $B \in A$

$$B = \Omega \cap B = \bigcup_{i=1}^{n} (A_i \cap B)$$

Dan karena $A_i \cap A_j = \varnothing$, $i \neq j$ sehingga

$$(A_i \cap B) \cap (A_j \cap B) = (A_i \cap A_j \cap B) = \varnothing \cap B = \varnothing$$

maka masing-masing peristiwa $(A_i \cap B)$ adalah saling asing.

Dari sini

$$P(B) = P\left[\bigcup_{i=1}^{n} (A_i \cap B)\right]$$

$$= \sum_{i=1}^{n} P(A_i \cap B)$$

$$= \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

Teorema 2. Aturan Bayes.

Diberikan suatu ruang probabilitas $\{\Omega, A, P\}$. Jika $A_1, A_2, ..., A_n$ merupakan suatu koleksi himpunan peristiwa saling asing dalam A yang memenuhi $\Omega = \bigcup_{i=1}^{n} A_i$ dan $P(A_i) > 0$ untuk i = 1, 2,........, n, maka untuk

sebarang peristiwa $B \in A$ dimana $P(B) > 0$. Maka untuk k = 1,..., n,

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

Bukti :

Karena

$$P(A_k \cap B) = P(B|A_k)P(A_k)$$

dan

$$P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

maka

$$P(A_k|B) = \frac{P(A_k \cap B)}{P(B)}$$

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^{n} P(B|A_i)P(A_i)}$$

2.2. Variabel Random (Peubah Acak)

Definisi 5. Variabel Random.

(Ω, A, P) adalah ruang probabilitas. X fungsi berharga riil pada Ω yang memenuhi setiap interval R

$$\{X \in R\} = \{\omega \in \Omega | X(\omega) \in R\} \in A$$

X disebut variabel random pada (Ω, A, P)

Definisi 6. Variabel Random Diskrit.

Jika setiap interval real memuat finite (countable) titik anggota A maka X disebut variabel random diskrit.

Definisi 7. Variabel Random Kontinu.

Jika setiap interval real memuat uncountable titik anggota A maka X disebut variabel random kontinu.

Definisi 8. Fungsi Kepadatan Probabilitas.

Fungsi f(x) adalah fungsi kepadatan probabilitas variabel random kontinu X yang didefinisikan di atas himpunan semua bilangan riil R, bila

1. $f(x) \geq 0$, untuk semua $x \in R$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$
3. $P(a < x < b) = \int_{a}^{b} f(x)dx$

Definisi 9. Distribusi Komulatif.

Distribusi kumulatif F(x) suatu variabel random kontinu X dengan fungsi kepadatan f(x) diberikan oleh :

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)dt$$

Definisi 10. Mean Variabel Random

Mean variabel random dinyatakan dengan μ, dimana

$$\mu = \sum_i x_i f(x_i), \text{ untuk X bernilai diskrit}$$

$$= \int_{-\infty}^{\infty} x f(x)dx, \text{ untuk X bernilai kontinu}$$

Definisi 11. Variansi.

Variasi didefinisikan sebagai berikut :

$$\sigma^2 = \sum_i (x_i - \mu)^2 f(x_i) \text{ untuk X bernilai diskrit}$$

$$= \int_{-\infty}^{\infty} (x_i - \mu)^2 f(x)dx \text{ untuk X bernilai kontinu}$$

atau

$$\sigma^2 = \sum x_i^2 f(x_i) - \mu^2 \text{ untuk X bernilai diskrit}$$

$$= \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2 \text{ untuk X bernilai kontinu}$$

2.3. Distribusi Normal

Definisi 12. Distribusi Normal

Sebuah variabel random X, disebut mempunyai sebuah distribusi normal jika mempunyai fungsi kepadatan :

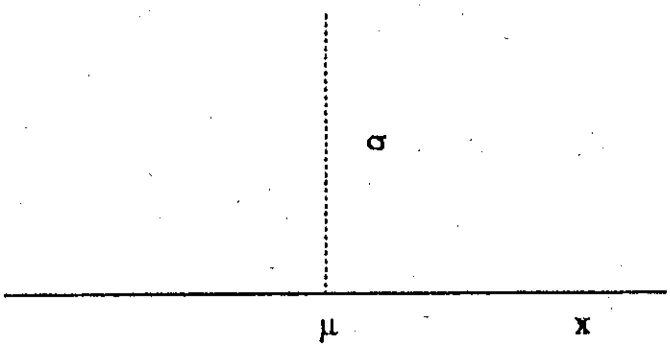
$$f(x) = N\left(\mu,\sigma^2\right) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\left(\frac{1}{2}\right)\left[\frac{x-\mu}{\sigma}\right]^2}, -\sim < x < \sim$$

$$0 < \sigma^2 < \infty$$

dengan mean μ dan variansi σ^2

dan $\pi = 3{,}14159\ldots$, $e = 2{,}71828\ldots$

Grafik fungsi kepadatan distribusi normal digambarkan sebagai berikut ;



Gambar 1. Kurva normal

Dari gambar 1 dapat diperoleh sifat-sifat distribusi normal :

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

2. $f(x) \geq 0$ untuk seluruh x

3. $\lim_{x\to\infty} f(x) = 0$ dan $\lim_{x\to-\infty} f(x) = 0$

4. Nilai maksimum dari f terjadi pada $x = \mu$

5. Titik belok dari f adalah $x = \mu \pm \sigma$, cekung dari bawah jika $\mu - \sigma < x < \mu + \sigma$ dan cekung atas untuk harga x lainnya.

Definisi 13. Distribusi Normal Standart.

Distribusi normal suatu variabel random dengan mean nol dan variasi satu disebut distribusi normal standart.

Definisi 14. Interval Kepercayaan.

Interval Kepercayaan (1 - α) 100% untuk μ ialah :

$$\overline{X} - z_{\frac{\alpha}{2}} . \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{\frac{\alpha}{2}} . \frac{\sigma}{\sqrt{n}}$$

dengan X menyatakan mean sampel ukuran n dari populasi dengan variansi σ^2 yang diketahui dan $z_{\frac{\alpha}{2}}$ menyatakan nilai distribusi normal baku sehingga daerah di sebelah kanannya memiliki luas $\frac{\alpha}{2}$

Contoh.

Mean dari simpangan baku nilai matematika sampel random 36 mahasiswa semester VI masing-masing 2,6 dan 0,3. Sehingga estimasi titik untuk μ ialah x = 2,6, dan karena ukuran sampelnya besar maka simpangan baku dapat dihampiri dengan s = 0,3.

$$z_{\frac{\alpha}{2}} = z_{\frac{0,05}{2}}$$

$z_{0,025} = 1,96$ (tabel).

Interval kepercayaan 95% untuk μ adalah :

$$2,6 - (1,96)\left(\frac{0,3}{\sqrt{36}}\right) < \mu < 2,6 + (1,96)\left(\frac{0,3}{\sqrt{36}}\right)$$

$$2,6 - 0,098 < \mu < 2,6 + 0,098$$

$$2,502 < \mu < 2,698$$

Definisi 15. Vektor Random.

Vektor random adalah vektor yang elemen-elemennya variabel random.

Definisi 16. Matriks Random.

Matriks random adalah matriks yang elemen-elemennya variabel random.

Definisi 17. Mean vektor random

Mean vektor random x dapat ditulis sebagai matriks yaitu

$$\mathrm{E}(x) = \begin{bmatrix} E(x_1) \\ E(x_{21}) \\ \vdots \\ E(x_p) \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix} = \mu$$

dengan

$$\mu_i = E(x_i) = \sum x_i p(x_i)$$

bila x_i variabel random diskrit dengan fungsi probabiltas $p_i(x_i)$

$$= \int_{-\infty}^{\infty} x_i f(x_i) dx_i$$

bila x_i variabel random kontinu dengan fungsi densitas $f(x_i)$

Definisi 18. Kovariant Vektor Random.

$$\sum = E(x-\mu)(x-\mu)^1 = E\left(\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \\ \vdots \\ x_p-\mu_p \end{bmatrix}\left[x_1-\mu_1, x_2-\mu_2, \cdots, x_p-\mu_p\right]\right)$$

$$= \begin{bmatrix} (x_1-\mu_1)^2 & (x_1-\mu_1)(x_2-\mu_2) & \cdots & (x_1-\mu_1)(x_p-\mu_p) \\ (x_2-\mu_2)(x_1-\mu_1) & (x_2-\mu_2)(x_2-\mu_2) & \cdots & (x_2-\mu_2)(x_p-\mu_p) \\ \vdots & \vdots & & \vdots \\ (x_p-\mu_p)(x_1-\mu_1) & (x_p-\mu_p)(x_2-\mu_2) & \cdots & (x_p-\mu_p)^2 \end{bmatrix}$$

$$= \begin{bmatrix} E(x_1-\mu_1)^2 & E(x_1-\mu_1)(x_2-\mu_2) & \cdots & E(x_1-\mu_1)(x_p-\mu_p) \\ E(x_2-\mu_2)(x_1-\mu_1) & E(x_2-\mu_2)(x_2-\mu_2) & \cdots & E(x_2-\mu_2)(x_p-\mu_p) \\ \vdots & \vdots & & \vdots \\ E(x_p-\mu_p)(x_1-\mu_1) & E(x_p-\mu_p)(x_2-\mu_2) & \cdots & E(x_p-\mu_p)^2 \end{bmatrix}$$

$$\sum = \mathrm{cov}(x) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

Karena $\sigma_{ik} = \sigma_{ki}$ maka

$$\sum = \mathrm{cov}(x) = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{12} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & \vdots & & \vdots \\ \sigma_{1p} & \sigma_{2p} & \cdots & \sigma_{pp} \end{bmatrix}$$

Definisi 19. Distribusi Normal Multivariant

Dengan demikian distribusi normal multivariat p dimensi untuk vektor random $\mathbf{x} = [x_1, x_2, \dots, x_p]^t$ mempunyai bentuk

$$f(x) = \frac{1}{(2\pi)^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)} \qquad -\sim < x_i < \sim,\ i = 1,2,\dots,p$$

diberi notasi $Np\ (\mu, \Sigma)$

2.4. Estimasi

Definisi 20. Inferensi Statistik.

Inferensi statistik adalah proses memperoleh informasi dari data sampel yang digunakan untuk menarik kesimpulan tentang populasi dari sampel yang dipilih.

Definisi 21. Fungsi Kepadatan Probabilitas.

X adalah sebuah variabel random dengan fungsi probabilitas f(x), dengan parameter θ yang tidak diketahui, maka fungsi kepadatan probabilitasnya adalah $f(x, \theta)$, $\theta \in \Omega$, dengan θ = ruang parameter.

Contoh.

Fungsi kepadatan probabilitas variabel random normal X, bergantung pada dua parameter dan dinyatakan dengan $n(x;\ \mu,\ \sigma)$. Kesimpulan distribusi normal $N(\theta,1)$ dinyatakan dengan $(N(\theta,1),\ \theta \in \Omega)$; $-\sim < \theta < \sim$.

Anggota dari kumpulan distribusi ini ialah N(0,1) merupakan distribusi normal standart.

Definisi 22. Distribusi Prior.

Misalkan $X_1, X_2, \ldots, X_n$, adalah suatu sampel random dari X, dan θ estimasi titik suatu parameter populasi θ yang tidak diketahui, maka $f(\theta)$ disebut distribusi prior.

Definisi 23. Fungsi Likelihood.

Misal X variabel random dengan fungsi probabilitas f(x) dengan parameter tidak diketahui dan $X_1 = x_1$, $X_2 = x_2$, ... , $X_n = x_n$ adalah data sampel, maka fungsi likelihood sampel adalah :

$$L(\theta) = f(x_1; \theta), f(x_2; \theta), \ldots, f(x_n; \theta), \theta \in \Omega$$

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

Definisi 24. Fungsi Kepadatan Marginal

Fungsi kepadatan marginal sampel $X_1, X_2, \ldots, X_n$ dengan parameter θ adalah :

$$u(x_1, x_2, \ldots, x_n) = \sum_{\theta} f(x_1, x_2, \ldots, x_n; \theta), \qquad \text{diskrit}$$

$$= \int_{-\infty}^{\infty} f(x_1, x_2, \ldots, x_n; \theta) d\theta \qquad \text{kontinu}$$

Definisi 25. Fungsi Kepadatan Bersama.

Fungsi kepadatan bersama sampel X_1, X_2,..., X_n dengan parameter θ adalah :

$$f(x_1, x_2, \ldots, x_n; \theta) = f(x_1, x_2, \ldots, x_n | \theta) f(\theta)$$

Definisi 26. Fungsi Kepadatan Bersyarat.

Fungsi kepadatan bersyarat sampel X_1, X_2,..., X_n dengan parameter θ adalah :

$$f(\theta | x_1, x_2, \ldots, x_n) = \frac{f(x_1, x_2, \ldots, x_n | \theta) f(\theta)}{u(x_1, x_2, \ldots, x_n)}$$

dapat ditulis : $f(\theta | y) = \dfrac{f(\theta, y)}{f(y)}$

Dengan $Y = y = u(x_1, x_2, \ldots, x_n)$ dan $f(\theta | y)$ disebut distribusi posterior.

2.6. Konsep Dasar Metode Peramalan

Data yang direkam dalam interval waktu yang sama dalam jangka waktu yang relatif panjang disebut data runtun waktu.

Interval waktu perekaman dapat amat singkat dan dapat cukup panjang , tergantung dari macamnya data yang direkam. Analisa runtun waktu umumnya memerlukan cacah data yang banyak. Oleh karena itu diperlukan rekaman yang panjang.

Data runtun waktu bervariasi karena adanya komponen-komponen trend, siklis, musiman dan komponen yang tidak teratur di dalamnya. Kenaikan permintaan beras yang disebabkan oleh laju kenaikan jumlah penduduk yang tetap besarnya adalah tergolong pengaruh trend. Naik turunnya curah hujan harian dalam kurun waktu beberapa tahun mengandung pengaruh musiman. Era kemakmuran mengandung komponen siklis. Terhambatnya produksi tekstil selama satu bulan karena kebakaran pabrik, mengandung komponen tidak teratur.

Dalam membuat forecast keadaan sosial pada umumnya dan bidang ekonomi pada kusunya tidak mungkin bisa tepat. Penyimpangan pasti ada karena tingkah laku manusia itu selalu dipengaruhi oleh berbagai macam hal, seperti kebudayaan, selera, perasaan dan sebagainya. Dalam bidang sosial dan ekonomi, meskipun kita tidak bisa membuat forecast yang persis sama dengan kenyataan, tetapi bukan berarti forecast ini tidak penting. Forecast sangat penting sebagai pedoman dalam pembuatan rencana. Kerja menggunakan forecast akan jauh lebih baik daripada tanpa forecast sama sekali. Hanya permasalahannya bagaimana membuat forecast agar bisa mendekati kenyataan. Untuk itu harus bisa memilih metode forecast yang paling cocok dengan masalahnya. Banyak sekali metode forecasting yang ada misalnya metode moving averages, metode eksponential smoothing, metode regresi dan sebagainya. Tidak ada metode forecast yang paling baik dan selalu cocok digunakan untuk membuat forecast setiap macam hal. Suatu metode mungkin cocok untuk membuat forecast mengenai sesuatu hal tetapi tidak cocok untuk membuat forecast hal yang lain. Untuk itu kita harus memilih metode yang cocok yaitu yang bisa meminimalkan forecast.

Dalam metode peramalan kita dapat menemukan model-model penyajian data runtun waktu, seperti model konstan, model untuk proses trend linier, model kuadrat, model regresi dan lain sebagainya. Beberapa model yang biasanya disajikan dalam bentuk runtun waktu adalah algebra atau fungsi dari waktu atau kombinasi dari keduanya. Sebagai contoh, jika observasi adalah sampel random dari beberapa distribusi probabilitas dengan mean dari distribusi tidak berubah karena waktu maka bisa digunakan model konstan.

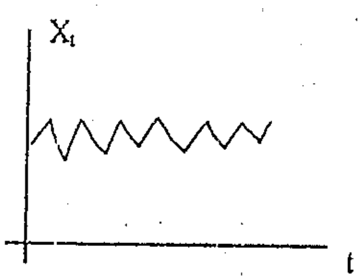
$$X_t = b + \varepsilon_t$$

Dimana X_t adalah permintaan pada periode t.

b adalah mean yang tidak diketahui

ε_t adalah random error dengan $E(\varepsilon_t) = 0$ dan varian $\varepsilon_t = \sigma_\varepsilon^2$.

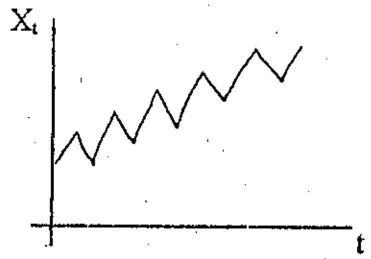
Untuk kasus konstan dapat digambarkan :



Gambar 2. Model Konstan

Untuk menyajikan suatu proses dari data runtun waktu jika mean berubah secara linier dari waktu maka digunakan model trend linier.

$$X_t = b_1 + b_2 t + \varepsilon_t$$



Gambar 3. Model Trend Linier

Sedangkan untuk data runtun waktu yang mempunyai kecenderungan berubah secara kuadratik dengan waktu maka model yang digunakan

$$X_t = b_1 + b_2 t + b_3 t^2 + \varepsilon_t$$

Sedangkan penyajian data runtun waktu yang berubah karena komponen siklis dapat disajikan misalnya :

$$X_t = b_1 + b_2 \sin \frac{2\pi.t}{12} + b_3 \cos \frac{2\pi.t}{12} + \varepsilon_t$$