BAB II

RUANG PROBABILITAS

2.1. *Probabilitas*

Probabilitas pada umumnya digunakan untuk mengukur kemungkinan terjadinya suatu peristiwa. Berkaitan dengan hal-hal tersebut, akan diuraikan beberapa definisi dan sifat-sifat dari probabilitas yang menunjang pengertian-pengertian dari statistik non-parametrik.

Definisi 2.1.1 Ruang sampel

Ruang sampel (Ω) adalah koleksi semua hasil percobaan yang mungkin terjadi.

Definisi 2.1.2 Kejadian dan Ruang kejadian

Kejadian adalah himpunan bagian dari Ruang sampel.

Ruang kejadian $(\mathcal{A})$ adalah kelas dari semua kejadian-kejadian yang berhubungan dengan suatu percobaan.

Definisi 2.1.3 Kejadian-kejadian yang saling asing

Himpunan bagian A dan B dari Ω didefinisikan sebagai *saling asing* jika $A \cap B = \emptyset$ atau $AB = \emptyset$. Secara lebih umum, himpunan-himpunan bagian $A_1 A_2, \ldots$ dari Ω didefinisikan sebagai *saling asing* jika

$$A_i \cap A_j = \emptyset \quad \text{atau} \quad A_i A_j = \emptyset$$

Berikut akan diuraikan dasar-dasar dari fungsi dan fungsi probabilitas.

Definisi 2.1.4 Fungsi

Suatu *fungsi*, misal $f(\cdot)$, dengan domain $\mathbf{A}$ dan kodomain $\mathbf{B}$ adalah koleksi pasangan berurut, misal (a,b) yang memenuhi

(i) $a \in \mathbf{A}$ dan $b \in \mathbf{B}$

(ii) $\forall\ a \in \mathbf{A}$ menjadi elemen pertama dari pasangan berurutan dalam koleksi

(iii) tidak ada dua pasangan berurutan (yang berbeda) dalam koleksi dengan elemen pertama yang sama

Definisi 2.1.5 Fungsi Probabilitas

Fungsi Probabilitas $P[\cdot]$ adalah himpunan fungsi dengan domain $\mathscr{A}$ (suatu aljabar dari kejadian-kejadian) dan kodomain interval [0,1] yang memenuhi

(i) $P[A] \geq 0$ untuk $\forall\ A \in \mathscr{A}$

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

(iii) Jika $A_1, A_2, \ldots$ adalah barisan yang saling asing dari kejadian-kejadian dalam $\mathscr{A}$ (yaitu, $A_i \cap A_j = \emptyset$ untuk $i \neq j; i,j=1,2,\ldots$) dan jika $A_1 \cup A_2 \cup \ldots = \bigcup_{i=1}^{\infty} A_i \in \mathscr{A}$,maka

$$P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[A_i]$$

Dari definisi-definisi di atas, maka dapat didefinisikan suatu Ruang Probabilitas, yaitu

Definisi 2.1.6 Ruang Probabilitas

Ruang Probabilitas adalah triplet $(\Omega, \mathscr{A}, P[\cdot])$, dimana Ω adalah ruang sampel, $\mathscr{A}$ adalah koleksi (diasumsikan suatu aljabar) kejadian-kejadian (masing-masing himpunan bagian dari Ω), dan $P[\cdot]$ adalah fungsi probabilitas dengan domain $\mathscr{A}$.

Definisi 2.1.7 Kejadian-kejadian yang saling bebas (independen)

Untuk suatu ruang probabilitas $(\Omega, \mathscr{A}, P[\cdot])$, diambil A dan B yang merupakan dua kejadian dalam $\mathscr{A}$. Kejadian-kejadian A dan B didefinisikan sebagai saling bebas (independen) jika dan hanya jika dipenuhi :

$$P[\,A\ B\,] = P[A]\ P[B]$$

Teorema 2.1.1 Penjumlahan dari probabilitas

Jika $A_1, A_2, \ldots, A_n$ adalah kejadian-kejadian yang saling asing dalam $\mathscr{A}$, maka

$$P[A_1 \cup A_2 \cup \cdots \cup A_n] = \sum_{j=1}^{n} P[A_j]$$

Bukti

Diambil $A_{n+1}, A_{n+2}, \ldots = \emptyset$, maka

$$\bigcup_{j=1}^{n} A_j = \bigcup_{j=1}^{\infty} A_j \in \mathscr{A}$$

dan dari definisi dipenuhi

$$P[\bigcup_{j=1}^{n} A_j] = P[\bigcup_{j=1}^{\infty} A_j] = \sum_{j=1}^{\infty} P[A_j] = \sum_{j=1}^{n} P[A_j]$$

Sehingga terbukti teorema di atas.

■

Untuk kejadian-kejadian dalam suatu ruang probabilitas, maka secara umum didapat aksioma-aksioma berikut

Aksioma I

Kejadian-kejadian pada $\mathscr{A}$, yaitu A^c, $\bigcap_j A_j$, dan $\bigcup_j A_j$ adalah juga merupakan kejadian.

Aksioma II

Probabilitas P pada $\mathscr{A}$ memenuhi $P\Omega = 1$, $PA \geq 0$, dan $P\sum_j A_j = \sum_j P\, A_j$.

Teorema 2.1.2 Ketaksamaan Boole

Jika $A_1, A_2, \ldots, A_n \in \mathscr{A}$, maka

$$P[A_1 \cup A_2 \cup \cdots \cup A_n] \leq P[A_1] + P[A_2] + \ldots + P[A_n]$$

Bukti

Untuk setiap dua kejadian A dan B anggota $\mathscr{A}$, diambil

$$A \cup B = A \cup A^c B \text{ dan } A \cap A^c B = \emptyset$$

sehingga berdasarkan *Teorema 2.1.1* akan dipenuhi

$$P[A \cup B] = P[A] + P[A^c B]$$

Kemudian diambil $B = AB \cup A^c B$ dan $AB \cap A^c B = \emptyset$, sehingga

$$P[B] = P[AB] + P[A^c B]$$

Dari persamaan di atas diperoleh

$$P[A^c B] = P[B] - P[AB]$$

sehingga untuk $A, B \in \mathscr{A}$, terbukti bahwa

$$\begin{aligned} P[A \cup B] &= P[A] + P[A^c B] \\ &= P[A] + P[B] - P[AB] \end{aligned}$$

$$\leq P[A] + P[B]$$

Secara umum , untuk n kejadian $\in \mathscr{A}$, didapat

$$P[\bigcup_{j=1}^{n} A_j] = \sum_{j=1}^{n} P[A_j] - \sum\sum_{i<j} P[A_i A_j] + \sum\sum\sum_{i<j<k} P[A_i A_j A_k]$$
$$- \ldots + (-1)^{n+1} P[A_1 A_2 \ldots A_n]$$
$$\leq \sum_{j=1}^{n} P[A_j]$$

Andaikan dipenuhi untuk n = k, maka berlaku

$$P[\bigcup_{j=1}^{k} A_j] = \sum_{j=1}^{k} P[A_j] - \sum\sum_{i<j} P[A_i A_j] + \sum\sum\sum_{i<j<k} P[A_i A_j A_k]$$
$$- \ldots + (-1)^{k+1} P[A_1 A_2 \ldots A_k]$$
$$\leq \sum_{j=1}^{k} P[A_j]$$

Maka untuk n = k + 1

$$P[\bigcup_{j=1}^{k+1} A_j] = \sum_{j=1}^{k+1} P[A_j] - \sum\sum_{i<j} P[A_i A_j] + \sum\sum\sum_{i<j<k} P[A_i A_j A_k]$$
$$- \ldots + (-1)^{k+1+1} P[A_1 A_2 \ldots A_{k+1}]$$
$$= (\sum_{j=1}^{k} P[A_j]) + P[A_{k+1}] - \sum\sum_{i<j} P[A_i A_j] + \sum\sum\sum_{i<j<k} P[A_i A_j A_k]$$
$$- \ldots + (-1)^{k+1+1} P[A_1 A_2 \ldots A_k A_{k+1}]$$
$$\leq (\sum_{j=1}^{k} P[A_j]) + P[A_{k+1}]$$
$$\leq \sum_{j=1}^{k+1} P[A_j]$$

Jadi terbukti teorema di atas.

■

2.2. *Variabel Acak dan Fungsi Distribusi*

Pada bagian ini akan dibicarakan konsep-konsep mengenai *variabel acak*, *fungsi distribusi kumulatif* dan *fungsi densitas* yang berkaitan dengan statistik non-parametrik. *Variabel acak* digunakan untuk menggambarkan kejadian-kejadian dalam suatu percobaan, sedangkan *fungsi distribusi kumulatif* digunakan untuk memberikan probabilitas dari kejadian-kejadian tertentu didefinisikan dalam kaitannya dengan variabel acak atau menggambarkan distribusi nilai-nilai dari variabel acak. Sedangkan untuk dua jenis variabel acak yang berbeda, yaitu diskret dan kontinu, dapat didefinisikan suatu *fungsi densitas* yang juga menggambarkan distribusi dari nilai-nilai variabel acak.

Definisi 2.2.1 Variabel acak

Untuk suatu ruang probabilitas $(\Omega, \mathscr{A}, P[\cdot])$, suatu *variabel acak* ($\mathscr{X}$ atau $\mathscr{X}(\cdot)$) adalah suatu fungsi dengan domain Ω dan kodomain garis riil. Fungsi $\mathscr{X}(\cdot)$ pasti merupakan himpunan A_r, yang didefinisikan sebagai $A_r = \{\omega : \mathscr{X}(\omega) \leq r\} \subset \mathscr{A}$ untuk $\forall\ r \in \mathbb{R}$.

Beberapa sifat dari variabel acak adalah

Definisi 2.2.2 Quantil dari variabel acak

Bilangan x_p ,untuk $0 \leq p \leq 1$ disebut *kuantil ke* p *dari variabel acak* $\mathcal{X}$, jika

$$P(\mathcal{X} < x_p) \leq p \quad \text{dan} \quad P(\mathcal{X} > x_p) \leq 1 - p$$

Definisi 2.2.3 Nilai Harapan

$\mathcal{X}$ adalah variabel acak dengan fungsi probabilitas $f(x)$ dan $u(\mathcal{X})$ sebagai fungsi berharga riil dari $\mathcal{X}$. *Nilai harapan* dari $u(\mathcal{X})$ dinotasikan sebagai $E[u(\mathcal{X})]$ didefinisikan sebagai

$$E[u(\mathcal{X})] = \sum_x u(x)\, f(x)$$

dimana notasi sigma tersebut meliputi seluruh nilai-nilai yang mungkin dari $\mathcal{X}$.

Dari definisi di atas terlihat bahwa rata-rata dari suatu variabel acak $\mathcal{X}$ dengan fungsi probabilitas $f(x)$ adalah

$$\mu = E[\mathcal{X}] = \sum_x x\, f(x)$$

Definisi 2.2.4 Varian dari variabel acak

$\mathcal{X}$ adalah variabel acak dengan rata-rata μ dan fungsi probabilitas $f(x)$, maka *Varian dari* $\mathcal{X}$, dinotasikan dengan σ^2 atau $\mathrm{Var}(\mathcal{X})$ didefinisikan sebagai

$$\sigma^2 = E[(\mathcal{X} - \mu)^2]$$

Dari definisi di atas dan dengan menggunakan definisi dari nilai harapan, maka varian dari suatu variabel acak $\mathcal{X}$ dapat disajikan sebagai

$$\sigma^2 = \sum_x (x - \mu)^2 \, f(x)$$
$$= \sum_x (x^2 - 2\mu x + \mu^2) f(x)$$
$$= \sum_x x^2 f(x) - 2\mu \sum_x x f(x) + \mu^2 \sum_x f(x)$$

karena $\sum_x f(x) = 1$, dan karena $\mu = \sum_x x f(x)$, maka persamaan di atas menjadi

$$\sigma^2 = \mathrm{E}(\mathcal{X}^2) - 2\mu^2 + \mu^2 = \mathrm{E}(\mathcal{X}^2) - \mu^2$$

Teorema 2.2.1 Kesamaan Bienayme

Jika $\mathcal{X}_k$, k = 1,2,...,n, adalah variabel acak - variabel acak yang saling bebas satu sama lain (saling independen), maka

$$\sigma^2 \sum_{k=1}^{n} \mathcal{X}_k = \sum_{k=1}^{n} \sigma^2 \mathcal{X}_k$$

Bukti

Karena $\mathrm{E}(\mathcal{X}_k - \mathrm{E}\mathcal{X}_k) = \mathrm{E}\mathcal{X}_k - \mathrm{E}\mathcal{X}_k = 0$, dan independensi dari $\mathcal{X}_k$ mengakibatkan independensi dari $\mathcal{X}_k - \mathrm{E}\mathcal{X}_k$, sehingga

$$
\begin{aligned}
\sigma^2\sum_{k=1}^{n} \mathcal{X}_k &= E(\sum_{k=1}^{n} \mathcal{X}_k - \sum_{k=1}^{n} E\mathcal{X}_k)^2 \\
&= E\{\sum_{k=1}^{n} (\mathcal{X}_k - E\mathcal{X}_k)\}^2 \\
&= \sum_{k=1}^{n} E(\mathcal{X}_k - E\mathcal{X}_k)^2 + \sum_{j\neq k=1}^{n} E(\mathcal{X}_j - E\mathcal{X}_j)(\mathcal{X}_k - E\mathcal{X}_k) \\
&= \sum_{k=1}^{n} \sigma^2\mathcal{X}_k + \sum_{j\neq k=1}^{n} E(\mathcal{X}_j - E\mathcal{X}_j)E(\mathcal{X}_k - E\mathcal{X}_k) \\
&= \sum_{k=1}^{n} \sigma^2\mathcal{X}_k
\end{aligned}
$$

■

Untuk setiap variabel acak didefinisikan suatu fungsi distribusi kumulatif, yaitu

Definisi 2.2.5 Fungsi Distribusi Kumulatif

Fungsi Distribusi Kumulatif dari suatu variabel acak $\mathcal{X}$, disimbolkan dengan $F_{\mathcal{X}}(\cdot)$, didefinisikan sebagai fungsi dengan domain garis riil dan kodomain interval [0,1] yang memenuhi

$$F_{\mathcal{X}}(x) = P[\ \mathcal{X} \le x\ ] = P[\ \{\omega : \mathcal{X}(\omega) \le x\}\ ]$$

untuk $\forall\ x \in \mathbb{R}$.

Jika fungsi distribusi kumulatifnya diketahui maka dapat digunakan untuk mencari probabilitas-probabilitas dari kejadian-kejadian yang didefinisikan dalam hubungannya dengan variabel acak yang berkaitan. Selanjutnya akan dibicarakan mengenai *fungsi densitas*, namun sebelumnya akan didefinisikan terlebih dahulu mengenai variabel acak diskret dan kontinu.

Definisi 2.2.6 Variabel acak Diskret

Suatu variabel acak $\mathcal{X}$ akan didefinisikan sebagai *diskret* jika daerah hasil (range) dari $\mathcal{X}$ adalah terbilang. Jika suatu variabel acak $\mathcal{X}$ diskret maka fungsi distribusi kumulatif yang berkorespondensi ($F_{\mathcal{X}}(\cdot)$) akan didefinisikan juga sebagai diskret.

atau

Variabel acak $\mathcal{X}$ dikatakan *diskret* jika terdapat korespondensi satu-satu antara nilai-nilai yang mungkin dari $\mathcal{X}$ dengan beberapa atau semua dari bilangan-bilangan bulat positif.

Definisi 2.2.7 Variabel acak Kontinu

Suatu variabel acak $\mathcal{X}$ didefinisikan sebagai *kontinu* jika terdapat suatu fungsi $f_{\mathcal{X}}(\cdot)$ sedemikian hingga

$$F_{\mathcal{X}}(x) = \int_{-\infty}^{x} f_{\mathcal{X}}(u)du$$

untuk setiap bilangan riil x.

atau

Variabel acak $\mathcal{X}$ dikatakan *kontinu* jika tidak terdapat dua kuantil x_p dan $x_{p'}$ dari $\mathcal{X}$ yang sama keduanya, dimana p tidak sama dengan p'. Secara ekivalen, suatu variabel acak $\mathcal{X}$ dikatakan *kontinu* jika $P(\mathcal{X} \leq x) = P(\mathcal{X} < x)$ untuk semua nilai-nilai x.

Dari definisi mengenai jenis dari variabel acak di atas, akan didefinisikan fungsi densitas (atau sering disebut dengan fungsi probabilitas) dari kedua jenis variabel acak di atas, yaitu

Definisi 2.2.8 Fungsi densitas diskret dari variabel acak diskret

Jika $\mathcal{X}$ adalah variabel acak diskret dengan nilai-nilai yang berbeda, yaitu $x_1, x_2, \ldots, x_n, \ldots$ maka fungsi yang dinotasikan dengan $f_{\mathcal{X}}(\cdot)$ dan didefinisikan dengan

$$f_{\mathcal{X}}(x) = \begin{cases} P[\mathcal{X}=x_n] & \text{jika } x = x_j \text{ , } j = 1,2,\ldots,n,\ldots \\ 0 & \text{jika } x \neq x_j \end{cases}$$

akan didefinisikan sebagai *fungsi densitas diskret* dari variabel acak diskret $\mathcal{X}$.

Fungsi densitas diskret juga dapat didefinisikan tanpa mengacu pada suatu variabel acak, yaitu

Definisi 2.2.9 Fungsi densitas diskret

Sebarang fungsi $f(\cdot)$ dengan daerah asal (domain) garis riil dan daerah kawan (kodomain) interval [0,1], didefinisikan sebagai *fungsi densitas diskret*

Jika untuk suatu himpunan terbilang $x_1, x_2, \ldots, x_n, \ldots$ berlaku

(i) $f(x_j) > 0$ untuk $j = 1,2,\ldots$

(ii) $f(x) = 0$ untuk $x \neq x_j; j=1,2,\ldots$

(iii) $\sum f(x_j) = 1$,dimana penjumlahan meliputi titik-titik $x_1, x_2, \ldots, x_n, \ldots$.

Definisi 2.2.10 Probabilitas fungsi densitas dari variabel acak kontinu

Jika $\mathcal{X}$ adalah variabel acak kontinu, fungsi $f_{\mathcal{X}}(\cdot)$ dalam

$$F_{\mathcal{X}}(x) = \int_{-\infty}^{x} f_{\mathcal{X}}(u)du$$

disebut *probabilitas fungsi densitas* dari variabel acak kontinu $\mathcal{X}$.

Fungsi densitas kontinu dapat juga didefinisikan tanpa mengacu pada variabel acak kontinu, yaitu

Definisi 2.2.11 Probabilitas fungsi densitas

Sebarang fungsi $f(\cdot)$ dengan daerah asal (domain) garis riil dan daerah kawan (kodomain) $[0,\infty)$ didefinisikan menjadi suatu *probabilitas fungsi densitas* jika dan hanya jika

(i) $f(x) \geq 0$ untuk semua x

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

Definisi 2.2.12 Fungsi probabilitas gabungan

Fungsi probabilitas gabungan $f(x_1, x_2, \ldots, x_n)$ dari variabel acak $X_1, X_2, \ldots, X_n$ adalah probabilitas dari kejadian gabungan $X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n$, dan didefinisikan sebagai

$$f(x_1, x_2, \ldots, x_n) = P(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

Definisi 2.2.13 Fungsi densitas saling bebas (independen)

$X_1, X_2, \ldots, X_n$ merupakan variabel acak berukuran n dengan fungsi densitas gabungan $f_{X_1, X_2, \ldots, X_n}(\cdot, \ldots, \cdot)$. Maka $X_1, X_2, \ldots, X_n$ dikatakan *saling bebas (independen)* jika dan hanya jika

$$f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$$

untuk semua nilai $x_1, x_2, \ldots, x_n$.

Definisi di atas berlaku untuk variabel acak diskret maupun variabel acak kontinu.

Definisi 2.2.14 Statistik

Statistik adalah fungsi dari beberapa variabel acak.

Definisi 2.2.15 Statistik Berurut

$x_1, x_2, \ldots, x_n$ adalah sampel-sampel acak berukuran n dari suatu fungsi distribusi kumulatif $F(\cdot)$. Maka

$$y_1 \le y_2 \le \cdots \le y_n$$

(dengan y_i merupakan nilai-nilai x_i yang diurutkan berdasarkan besarnya) didefinisikan sebagai *statistik berurut* yang berkorespondensi dengan sampel-sampel acak $x_1, x_2, \ldots, x_n$.

Statistik berurut memegang peranan yang penting pada penarikan kesimpulan suatu pengujian.

2.3. *Sampel Acak dan Fungsi Distribusi Kumulatif Sampel*

Pada eksperimen-eksperimen selalu didapatkan hasil-hasil yang terdiri dari data-data, sehingga berdasarkan data-data tersebut dapat ditarik suatu kesimpulan yang menyatakan hasil dari eksperimen tersebut. Guna mencapai hasil yang cukup akurat, maka eksperimen biasanya dilakukan dengan mengambil sampel-sampel dari suatu populasi, dan data-data dari sampel inilah yang akan diolah untuk menarik kesimpulan-kesimpulan dari suatu populasi. Untuk itu diperlukan suatu populasi sasaran yang akan diuji. Berikut ini akan didefinisikan mengenai populasi sasaran , yaitu

Definisi 2.3.1 Populasi sasaran

Seluruh anggota yang dibicarakan dan yang akan diambil kesimpulannya disebut *populasi sasaran*.

Dari populasi sasaran itu kemudian ditarik sampel acak yang akan digunakan untuk menganalisa populasi tersebut.

Definisi 2.3.2 Sampel acak

Variabel acak $X_1, X_2, \ldots, X_n$ yang memiliki fungsi densitas gabungan $f_{X_1, X_2, \ldots, X_n}(\cdot, \ldots, \cdot)$ dimana faktornya adalah

$$f_{X_1, X_2, \ldots, X_n}(x_1, x_2, \ldots, x_n) = f(x_1)f(x_2)\ldots f(x_n)$$

dengan $f(\cdot)$ merupakan fungsi densitas setiap X_i. Maka $X_1, X_2, \ldots, X_n$ merupakan *sampel acak* ukuran n dari suatu populasidengan fungsi densitas $f(\cdot)$.

Sampel acak dari suatu populasi dapat dianggap suatu populasi tersendiri, yaitu

Definisi 2.3.3 Populasi dimana sampel diambil

$X_1, X_2, \ldots, X_n$ sebagai sampel acak dari suatu populasi dengan fungsi densitas $f(\cdot)$, maka populasi ini disebut dengan *populasi dimana sampel diambil.*

Fungsi distribusi kumulatif dari sampel dapat digunakan untuk mengestimasi fungsi distribusi kumulatif dari populasi dimana sampel diambil, dengan anggapan fungsi distribusi kumulatif dari populasi tersebut tidak diketahui bentuk fungsinya atau paling sedikit sebagian tidak diketahui bentuk fungsinya.

Definisi 2.3.4 Fungsi Distribusi Kumulatif dari Sampel

$x_1, x_2, \ldots, x_n$ menyatakan suatu sampel acak dari suatu fungsi distribusi kumulatif $F(\cdot)$, dan

$$y_1 \le y_2 \le \cdots \le y_n$$

menyatakan statistik berurut yang berkorespondensi. *Fungsi distribusi kumulatif sampel* $(F_n(x))$ didefinisikan sebagai

$$F_n(x) = \left(\frac{1}{n}\right) \times (\text{banyaknya } y_j \text{ yang} \le x)$$

atau dengan sama dinyatakan dengan

$$F_n(x) = \left(\frac{1}{n}\right) \times (\text{banyaknya } x_i \text{ yang} \le x)$$

2.4. *Beberapa definisi penunjang*

Pada bagian ini akan diuraikan beberapa pengertian-pengertian yang berkaitan dengan analisa dari fungsi distribusi kumulatif dari sampel untuk membuktikan kekonvergenannya dengan fungsi distribusi kumulatif dari populasi sebagai dasar pemikiran dari statistik penguji Kolmogorov-Smirnov.

Definisi 2.4.1 Fungsi indikator

Ω adalah sebarang ruang dengan titik-titik ω dan A adalah sebarang himpunan bagian dari Ω. *Fungsi indikator* dari A, dinotasikan dengan $I_A(\cdot)$ adalah fungsi dengan daerah asal (domain) Ω dan daerah kawan (kodomain) himpunan yang terdiri dari dua bilangan real 0 dan 1 dan didefinisikan dengan

$$I_A(\omega) = \begin{cases} 1 & \text{jika } \omega \in A \\ 0 & \text{jika } \omega \notin A \end{cases}$$

Teorema 2.4.1 Ketaksamaan segitiga

Untuk sebarang a dan b dalam $\mathbb{R}$, maka berlaku

$$|a + b| \leq |a| + |b|$$

Bukti

Diandaikan $|a| \leq c$,dengan $c > 0$, maka sesuai definisi dari harga mutlak, didapat $a \leq c$ dan $-a \leq c$

($= -c \le a$) sehingga $-c \le a \le c$. Jika diambil $c = |a|$ maka berlakulah $-|a| \le a \le |a|$ dan $-|b| \le b \le |b|$. Karena untuk a, b, c dan d dalam $\mathbf{R}$ memenuhi untuk $a>b$ dan $c>d$ berlaku $a+c>b+d$, sehingga diperoleh hubungan

$$-(|a| + |b|) \le a + b \le |a| + |b|$$

oleh sebab itu berlaku

$$|a + b| \le |a| + |b|$$

■

Definisi 2.4.2 Sekitar

Diambil $a \in \mathbf{R}$. Maka

(i) untuk $\varepsilon > 0$, *sekitar-ε* dari a adalah himpunan $V_\varepsilon(a) := \{x \in \mathrm{R} : |x - a| < \varepsilon\}$

(ii) Suatu *sekitar* dari a adalah sebarang himpunan yang memuat suatu *sekitar-ε* dari a untuk suatu $\varepsilon > 0$.

Definisi 2.4.3 Batas atas

Diambil S sebagai himpunan bagian dari $\mathbf{R}$ (=himpunan bilangan riil). Suatu elemen $u \in \mathbf{R}$ dikatakan sebagai *batas atas* dari S jika $s \le u$ untuk semua $s \in S$.

Definisi 2.4.4 Supremum (batas atas terkecil)

Diambil S sebagai suatu himpunan bagian dari $\mathbf{R}$ (= himpunan bilangan riil). Jika S terbatas ke atas, maka suatu batas atas dikatakan sebagai *supremum* (atau *batas atas terkecil*) dari S, jika batas atas tersebut lebih kecil dari setiap batas atas yang lain dari S.

atau

$u \in \mathbb{R}$ adalah *supremum (batas atas terkecil)* dari $S \subseteq \mathbb{R}$ jika memenuhi dua syarat berikut :

(i) $s \leq u$ untuk semua $s \in S$;

(ii) jika $s \leq v$ untuk semua $s \in S$, maka $u \leq v$.

Definisi 2.4.5 Barisan konvergen

$\mathcal{X} = (x_n)$ adalah barisan bilangan riil. Bilangan riil x dikatakan *limit dari* $\mathcal{X}$ jika untuk setiap sekitar V dari x terdapat suatu bilangan asli $K(V)$ sedemikian sehingga untuk semua $n \geq K(V)$, x_n termasuk dalam V.

Teorema 2.4.2 Kekonvergenan barisan bilangan riil

$\mathcal{X} = (x_n)$ adalah barisan bilangan riil, dan $x \in \mathbb{R}$. Maka berlaku secara ekuivalen pernyataan-pernyataan berikut :

(a) $\mathcal{X}$ konvergen ke x

(b) Untuk setiap sekitar ε (V_ε) dari x terdapat suatu bilangan asli $K(\varepsilon)$ sedemikian sehingga untuk semua $n \geq K(\varepsilon)$, x_n termasuk dalam V_ε

(c) Untuk setiap $\varepsilon > 0$ terdapat suatu bilangan asli

$K(\varepsilon)$ sedemikian sehingga untuk semua $n \geq K(\varepsilon)$, maka $x - \varepsilon < x_n < x + \varepsilon$

(d) Untuk setiap $\varepsilon > 0$ terdapat suatu bilangan asli $K(\varepsilon)$ sedemikian sehingga untuk semua $n \geq K(\varepsilon)$, maka $|x_n - x| < \varepsilon$.

Bukti

(a) ⇒ (b)

Jika $\mathcal{X}$ konvergen ke x sesuai pernyataan *Definisi 2.4.5* maka karena sebuah sekitar ε, yaitu $V_\varepsilon = (x - \varepsilon , x + \varepsilon)$ dari x adalah suatu sekitar dari x, sehingga terdapat suatu bilangan asli $K(\varepsilon) := K(V_\varepsilon)$ sedimikian sehingga untuk semua $n \geq K(\varepsilon)$, berlaku x_n termasuk dalam V_ε.

(b) ⇒ (c)

$V_\varepsilon = (x - \varepsilon , x + \varepsilon)$ sehingga, jika $x_n \in V_\varepsilon$, maka $x - \varepsilon < x_n < x + \varepsilon$.

(c) ⇒ (d)

Jika $x - \varepsilon < x_n < x + \varepsilon$ berlaku, maka memenuhi

$$|x_n - x| < \varepsilon$$

(d) ⇒ (a)

Andaikan pernyataan (d) berlaku untuk sebarang pilihan $\varepsilon > 0$, maka akan diperlihatkan bahwa $x = (x_n)$ konvergen ke x sesuai dengan *Definisi 2.4.5.* Untuk itu diambil V sebagai sebarang sekitar dari x. Dengan *Definisi 2.4.2* , maka terdapat suatu $\varepsilon > 0$ sedemikian sehingga $V \supseteq V_\varepsilon = (x - \varepsilon , x + \varepsilon)$. Diambil $K(\varepsilon)$ sedemikian sehingga untuk semua $n \geq K(\varepsilon)$ maka $|x_n - x| < \varepsilon$. Dengan demikian berlaku juga untuk $n \geq K(\varepsilon)$, maka

$$x - \varepsilon < x_n < x + \varepsilon$$

Oleh sebab itu jika $n \geq K(\varepsilon)$, maka $x_n \in V$, maka $K(V)$ pada *Definisi 2.4.5* dapat diambil sama dengan $K(\varepsilon)$ untuk pilihan ε di atas.

■

Definisi 2.4.6 Barisan fungsi konvergen

(f_n) adalah barisan fungsi-fungsi pada himpunan $D \subseteq \mathbb{R}$ dengan daerah kawan (kodomain) $\mathbb{R}$ sendiri. Diambil $D_o \subseteq D$ dan f sebagai fungsi dengan daerah asal (domain) memuat D_o dan daerah hasil (range) pada $\mathbb{R}$. Barisan (f_n) dikatakan *konvergen* pada D_o ke f ,jika untuk setiap $x \in D_o$, barisan $(f_n(x))$ konvergen dalam $\mathbb{R}$ ke $f(x)$.

Definisi 2.4.7 Barisan fungsi konvergen seragam

Barisan fungsi-fungsi (f_n) pada $D \subseteq \mathbb{R}$ dengan daerah kawan (kodomain) $\mathbb{R}$ sendiri, dikatakan *konvergen seragam* pada suatu himpunan bagian D_o dari D pada fungsi f dalam kasus untuk setiap $\varepsilon > 0$,terdapat suatu bilangan asli $K(\varepsilon)$ (yang tergantung pada ε, tapi tidak pada $x \in D_o$) sedemikian sehingga untuk semua $n \geq K(\varepsilon)$ dan $x \in D_o$, maka

$$|f_n(x) - f(x)| < \varepsilon$$

Dua macam bentuk penaksiran merupakan pusat yang penting dalam statistika terapan. Bentuk pertama adalah, suatu variabel acak yang diberikan ditaksir dengan variabel acak yang lain. Dan bentuk kedua adalah, fungsi distribusi yang diberikan ditaksir dengan fungsi distribusi yang lain. Sehingga diperlukan jenis-jenis konvergensi yang dapat menggambarkan seberapa dekat penaksir tersebut dengan yang ditaksir. Berikut ini adalah definisi-definisi yang berkaitan dengan jenis-jenis konvergensi dari barisan variabel acak.

Definisi 2.4.8 Konvergen dalam probabilitas

$\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n$ dan $\mathcal{X}$ adalah variabel-variabel acak pada ruang probabilitas $(\Omega, \mathcal{A}, P[\cdot])$. $\mathcal{X}_n$ didefinisikan *konvergen dalam probabilitas* ke $\mathcal{X}$,jika

$$\lim_{n\to\infty} P(\ |\mathcal{X}_n - \mathcal{X}| < \varepsilon\ ) = 1$$

untuk setiap $\varepsilon > 0$.

atau

$\mathcal{X}_n \xrightarrow{P} \mathcal{X}$ jika untuk $\forall\ \varepsilon > 0$,maka

$$P\ [\ |\mathcal{X}_n - \mathcal{X}| \geq \varepsilon\ ] \longrightarrow 0$$

Definisi 2.4.9 Konvergen dengan probabilitas 1

$\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n$ dan $\mathcal{X}$ adalah variabel-variabel acak pada ruang probabilitas $(\Omega, \mathcal{A}, P[\cdot])$. $\mathcal{X}_n$ didefinisikan *konvergen dengan probabilitas 1* ke $\mathcal{X}$,jika

$$P\ (\ \lim_{n\to\infty} \mathcal{X}_n = \mathcal{X}\ ) = 1$$

atau

$\mathcal{X}_n \xrightarrow{vp\ 1} \mathcal{X}$ jika untuk $\forall\ \varepsilon > 0$, maka

$$P\ \bigcup_{k\geq n} [\ |\mathcal{X}_k - \mathcal{X}| \geq \varepsilon\ ] \longrightarrow 0$$

(vp 1 = dengan probabilitas 1)

Definisi 2.4.10 Konvergen dalam distribusi

$F_1(\cdot), F_2(\cdot), \ldots, F_n(\cdot)$ dan $F(\cdot)$ adalah fungsi-fungsi distribusi. $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_n$ dan $\mathcal{X}$ adalah variabel-variabel acak yang mempunyai distribusi-distribusi di atas secara berurutan. $\mathcal{X}_n$ didefinisikan *konvergen dalam distribusi* ke $\mathcal{X}$, jika

$$\lim_{n\to\infty} F_n(t) = F(t)$$

untuk setiap titik kontinuitas t dari F.

Berikut ini akan diuraikan mengenai kasus Bernoulli yang berkaitan dengan hukum bilangan besar dari

Bernoulli dan hukum kuat bilangan besar dari Borel. Suatu kejadian A_k, k = 1,2,...n, yang saling bebas satu sama lain (saling independen) sedemikian sehingga PA_k mempunyai nilai yang sama yaitu sama dengan p. Banyaknya hasil yang sukses dari kejadian A dalam n pengulangan dinyatakan dengan variabel acak $S_n = \sum_{k=1}^{n} I_{A_k}$ dengan I_{A_k} adalah fungsi indikator dari kejadian A_k. Untuk menyajikan S_n dalam bentuk biasa, yaitu dengan nilai-nilai yang dikawankan ke kejadian-kejadian dari himpunan bagian seluruh kejadian, maka I_{A_k} dapat dipandang sebagai

$$I_{A_k} = I_{A_k} \prod_{j=1 \& j \neq k}^{n} (I_{A_j} + I_{A_j^c})$$

dengan

$$\begin{aligned}\prod_{j=1 \& j \neq k}^{n} (I_{A_j} + I_{A_j^c}) &= (I_{A_1} + I_{A_1^c}) \ldots (I_{A_n} + I_{A_n^c}) \\ &= (I_{A_1} + 1 - I_{A_1}) \ldots (I_{A_n} + 1 - I_{A_n}) \\ &= (1)(1) \ldots (1) \\ &= 1\end{aligned}$$

Jika disubstitusikan ke S_n diperoleh

$$S_n = \sum_{j=0}^{n} j I_{B_j}$$

dengan $I_{B_j} = \sum I_{A_{k_1}} \ldots I_{A_{k_j}} I_{A^c_{k_{j+1}}} \ldots I_{A^c_{k_n}}$

Penjumlahan meliputi semua permutasi dari k = 1,2,...,n, yang dikelompokkan menjadi dua kelompok dengan j anggota dan n-j anggota. Karena kejadian-kejadian tersebut saling bebas satu sama lain, maka

$$PA_{k_1} PA_{k_2} \dots PA_{k_j} PA^c_{k_{j+1}} \dots PA^c_{k_n} = p^j q^{n-j}, \quad q = 1 - p$$

sehingga probabilitas terjadinya suatu kejadian sebanyak j kali pada n kali percobaan adalah

$$P[S_n = j] = P\ B_j = \binom{n}{j} p^j q^{n-j} \ ; \ j = 0,1,2,\dots,n$$

Dari hasil di atas didapatkan

$$ES_n = E\sum_{k=1}^{n} I_{A_k} = \sum_{k=1}^{n} PA_k = np$$

dan dengan kesamaan Bienayme untuk variabel-variabel acak yang saling bebas satu sama lain (saling independen) didapatkan

$$\begin{aligned}
\sigma^2 S_n = \sum_{k=1}^{n} \sigma^2 I_{A_k} &= \sum_{k=1}^{n} E(I_{A_k} - EI_{A_k})^2 \\
&= \sum_{k=1}^{n} EI^2_{A_k} - E^2 I_{A_k} \\
&= \sum_{k=1}^{n} EI_{A_k} - E^2 I_{A_k} \\
&= \sum_{k=1}^{n} (p - p^2) = npq
\end{aligned}$$

Teorema 2.4.3 Ketaksamaan Tchebichev

Jika $\mathcal{X}$ adalah variabel acak dengan definisi

$$\mathcal{X} = \sum_{j=1}^{m} x_j I_{A_j}$$

,maka untuk setiap $\varepsilon > 0$, berlaku

$$P[|X| \geq \varepsilon] \leq \frac{1}{\varepsilon^2} EX^2$$

dimana $[|X| \geq \varepsilon]$ adalah gabungan (union) dari semua kejadian dengan nilai $|X|$ adalah $\geq \varepsilon$.

Bukti

$$\begin{aligned} EX^2 &= E(X^2 I_{[|X| \geq \varepsilon]}) + E(X^2 I_{[|X| < \varepsilon]}) \\ &\geq E(X^2 I_{[|X| \geq \varepsilon]}) \\ &\geq \varepsilon^2 E I_{[|X| \geq \varepsilon]} = \varepsilon^2 P[|X| \geq \varepsilon] \end{aligned}$$

Sehingga terbukti

$$P[|X| \geq \varepsilon] \leq \frac{1}{\varepsilon^2} EX^2$$

■

Teorema 2.4.4 Kekonvergenan barisan variabel acak X_n *ke* 0

Jika untuk semua bilangan bulat m,

$$\sum_{n=1}^{\infty} P\left[|X_n| \geq \frac{1}{m}\right] < \infty$$

maka $P[X_n \not\longrightarrow 0] = 0$.

Bukti

Dibentuk kejadian $A_{nm} = \bigcup_{v=1}^{\infty} \left[|X_{n+v}| \geq \frac{1}{m}\right]$ dan kejadian $A_m = \bigcap_{n=1}^{\infty} A_{nm}$.

Dari ketaksamaan Boole dan dari hipotesa pernyataan di atas, maka untuk setiap m,

$$P\ A_{nm} = P \bigcup_{k=n+1}^{\infty} [|X_k| \geq \frac{1}{m}] \leq \sum_{k=n+1}^{\infty} P[|X_n| \geq \frac{1}{m}] \longrightarrow$$

0, pada n menuju ke ∞.

Karena untuk sebarang n´ berlaku

$$P\ A_m = P \bigcap_{n=1}^{\infty} A_{nm} \leq P\ A_{n'm}$$

sehingga jika diambil n´ menuju ke ∞ menyebabkan PA_m = 0. Maka dengan ketaksamaan Boole, dipenuhi

$$P[X_n \not\longrightarrow 0] = P \bigcup_{m=1}^{\infty} A_m \leq \sum_{m=1}^{\infty} PA_m = 0$$

Jadi terbukti teorema di atas.

∎

Teorema 2.4.5 Hukum bilangan besar Bernoulli (1713)

Dalam kasus Bernoulli untuk setiap $\varepsilon > 0$, pada n ⟶ ∞, maka

$$P\ [\ |\frac{S_n}{n} - p| \geq \varepsilon\ ] \longrightarrow 0$$

yaitu $\frac{S_n}{n}$ konvergen dalam probabilitas ke p pada n menuju ∞.

Bukti

Dengan ketaksamaan Tchebichev, didapat pada n menuju ∞, memenuhi

$$P\left[\,\left|\frac{S_n}{n}-p\right|\geq\varepsilon\,\right]=P\left[\,|S_n-ES_n|\geq\varepsilon n\,\right]$$
$$\leq\frac{1}{\varepsilon^2 n^2}\sigma^2 S_n=\frac{pq}{\varepsilon^2 n}\longrightarrow 0$$

Jadi terbukti $\frac{S_n}{n}$ konvergen dalam probabilitas ke p pada n menuju ∞.

■

Teorema 2.4.6 Hukum kuat bilangan besar Borel (1909)

Dalam kasus Bernoulli, maka

$$P\left[\,\frac{S_n}{n}\longrightarrow p\,\right]=1$$

yaitu $\frac{S_n}{n}$ konvergen dengan probabilitas 1 ke p.

Bukti

Dalam kasus Bernoulli $\mathcal{X}_n=\frac{S_n}{n}=\frac{1}{n}\sum_{j=1}^{n}I_{A_j}$

dimana A_j adalah kejadian-kejadian yang saling bebas satu sama lain dengan probabilitas yang sama yaitu p untuk sejumlah n kejadian.

$$E\mathcal{X}_n=E\left(\frac{S_n}{n}\right)=\frac{1}{n}E\sum_{j=1}^{n}I_{A_j}=\frac{1}{n}\sum_{j=1}^{n}PA_j=\frac{1}{n}np=p$$

dan

$$\sigma^2\mathcal{X}_n=\sigma^2\left(\frac{S_n}{n}\right)=\left(\frac{1}{n}\right)^2\sigma^2(S_n)$$
$$=\left(\frac{1}{n}\right)^2\sum_{j=1}^{n}\sigma^2 I_{A_j}$$
$$=\left(\frac{1}{n}\right)^2\sum_{j=1}^{n}[E(I_{A_j}-EI_{A_j})^2]$$

$$= \frac{1}{n} [EI^2_{A_j} - E^2 I_{A_j}]$$

$$= \frac{1}{n} [EI_{A_j} - E^2 I_{A_j}]$$

$$= \frac{1}{n} [p - p^2]$$

$$= \frac{1}{n} [p(1-p)] = \frac{1}{n} pq$$

Dari *Teorema 2.4.4* maka untuk setiap bilangan bulat m berlaku

$$\sum_{k=1}^{\infty} P [|\mathcal{X}_{k^2} - p| \geq \frac{1}{m}] = \sum_{k=1}^{\infty} P [|\frac{S_{k^2}}{k^2} - p| \geq \frac{1}{m}]$$

$$= \sum_{k=1}^{\infty} P [|S_{k^2} - ES_{k^2}| \geq \frac{k^2}{m}]$$

dan dari ketaksamaan Tchebichev didapat

$$\leq \sum_{k=1}^{\infty} \frac{m^2}{k^4} E(S_{k^2} - ES_{k^2})^2$$

$$\leq \sum_{k=1}^{\infty} \frac{m^2}{k^4} \sigma^2 S_{k^2}$$

$$\leq \sum_{k=1}^{\infty} \frac{m^2}{k^4} \sum_{j=1}^{k^2} \sigma^2 I_{A_j}$$

$$\leq \sum_{k=1}^{\infty} \frac{m^2}{k^4} k^2 pq$$

$$\leq \sum_{k=1}^{\infty} \frac{m^2}{k^2} pq$$

$$\leq m^2 pq \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$$

Jadi terlihat bahwa $P [\mathcal{X}_{k^2} \not\rightarrow p] = 0$, atau dengan kata lain $\mathcal{X}_{k^2} \longrightarrow p$ dengan probabilitas 1 pada k menuju ∞. Tetapi untuk $\forall$ n berkorespondensi dengan suatu bilangan bulat $k = k(n)$ dimana $k^2 \le n < (k+1)^2$, oleh sebab itu $0 \le n-k^2 \le 2k$ dan n menuju ∞ yang mengakibatkan k menuju ∞. Karena

$$|x_n - x_{k^2}| = \left| \frac{1}{n}\sum_{j=1}^{n} I_{A_j} - \frac{1}{k^2}\sum_{j=1}^{k^2} I_{A_j} \right|$$

$$= \left| \frac{1}{n}\sum_{j=1}^{k^2} I_{A_j} + \frac{1}{n}\sum_{j=k^2+1}^{n} I_{A_j} - \frac{1}{k^2}\sum_{j=1}^{k^2} I_{A_j} \right|$$

$$= \left| \left(\frac{1}{n} - \frac{1}{k^2}\right)\sum_{j=1}^{k^2} I_{A_j} + \frac{1}{n}\sum_{j=k^2+1}^{n} I_{A_j} \right|$$

dari ketaksamaan segitiga, maka

$$\le \left| \left(\frac{1}{n} - \frac{1}{k^2}\right)\sum_{j=1}^{k^2} I_{A_j} \right| + \left| \frac{1}{n}\sum_{j=k^2+1}^{n} I_{A_j} \right|$$

$$\le \left| \frac{k^2-n}{nk^2}(k^2) \right| + \left| \frac{1}{n}(n - k^2) \right|$$

$$\le \left| -\frac{(n-k^2)}{nk^2}(k^2) \right| + \left| \frac{1}{n}(n - k^2) \right|$$

$$\le \frac{(n-k^2)}{nk^2}(k^2) + \frac{1}{n}(n - k^2)$$

$$\le \frac{(n-k^2)}{n} + \frac{1}{n}(n - k^2)$$

$$\le 2\frac{(n-k^2)}{n} \le \frac{2(2k)}{n} \le \frac{4k}{n}$$

karena $k^2 \le n$, maka $\frac{1}{k^2} \ge \frac{1}{n}$, sehingga

$$\frac{4k}{n} \le \frac{4k}{k^2} \le \frac{4}{k}$$

Dengan demikian

$$\begin{aligned} |x_n - p| &= |(x_n - x_{k^2}) + (x_{k^2} - p)| \\ &\le |(x_n - x_{k^2})| + |(x_{k^2} - p)| \\ &\le \frac{4}{k} + |(x_{k^2} - p)| \end{aligned}$$

Karena $x_{k^2} \longrightarrow p$ dengan probabilitas 1, maka terlihat bahwa $x_n \longrightarrow p$ dengan probabilitas 1 pada n menuju ∞. Jadi terbukti hukum kuat bilangan besar dari Borel.

■