UNIVERSITAS GADJAH MADA
DEPARTEMEN TEKNIK SIPIL DAN LINGKUNGAN
PRODI MAGISTER TEKNIK PENGELOLAAN BENCANA ALAM

Teknik Pengolahan Data

Probabilitas

Probabilitas

- Probabilitas, Peluang, Kemungkinan
- Mengapa probabilitas?
  - Orang tidak dapat memastikan nilai suatu proses (misal erupsi gunung berapi) berdasarkan data erupsi selama waktu yang lalu sampai saat ini.
  - Sifat stokastik ataupun ketidak-pastian merupakan sifat yang melekat pada proses (yang melibatkan) alam.

Probabilitas, Peluang Kejadian

No urut	Jumlah hari terjadinya kemacetan pasokan air per bulan	Frekuensi
1	10	2
2	9	1
3	8	0
4	7	2
5	6	3
6	5	5
7	4	4
8	3	8
9	2	3
10	1	2
	Jumlah	30

Jumlah hari saat terjadi kemacetan pasokan air PDAM selama 30 bulan terakhir.

Dapatkah Saudara memastikan jumlah hari akan terjadi kemacetan pasokan air PDAM pada bulan depan?

Probabilitas, Peluang Kejadian, Risiko

Debit puncak suatu Sungai XYZ selama 66 tahun

Tahun ke-	Debit (m^3/s)	Tahun ke-	Debit (m^3/s)	Tahun ke-	Debit (m^3/s)
1	473	23	1110	45	843
2	544	24	717	46	450
3	872	25	961	47	284
4	657	26	925	48	460
5	915	27	341	49	804
6	535	28	690	50	550
7	678	29	734	51	729
8	700	30	991	52	712
9	669	31	792	53	468
10	347	32	626	54	841
11	580	33	937	55	613
12	470	34	687	56	871
13	663	35	801	57	705
14	809	36	323	58	777
15	800	37	431	59	442
16	523	38	770	60	206
17	580	39	536	61	850
18	672	40	708	62	829
19	115	41	894	63	887
20	461	42	626	64	602
21	524	43	1120	65	403
22	943	44	440	66	505

Dapatkah Saudara memastikan debit maksimum pada tahun ke-67?

Probabilitas

- Definisi #1
  - Andaikata suatu peristiwa random dapat terjadi dalam *n* cara yang masing-masing memiliki kemungkinan yang sama, dan apabila sejumlah n_a cara memberikan hasil *A*, maka probabilitas terjadinya peristiwa dengan hasil *A* adalah n_a/n

$$\text{prob}(A) = \frac{n_a}{n}$$

  - Dalam definisi di atas, *n* adalah himpunan semua yang mungkin terjadi.
  - Definisi di atas berasumsi bahwa *n* diketahui, padahal himpunan semua cara yang mungkin pada kenyataannya tidak selalu diketahui atau tidak terjadi atau tidak diamati atau tidak dihitung.

Probabilitas

- Definisi #2
  - Andaikata suatu peristiwa random terjadi berkali-kali dalam jumlah yang sangat besar, *n* kali, dan sejumlah n_a kali memiliki hasil *A*, maka probabilitas peristiwa dengan hasil *A* adalah

$$\mathrm{prob}(A) = \lim_{n \to \infty} \frac{n_a}{n}$$

  - Definisi di atas berbeda dengan definisi #1 dalam hal-hal berikut:
    - Probabilitas suatu kejadian "diperkirakan" (*can be estimated*) berdasarkan observasi sejumlah *n* kali.
    - *n* di sini tidak/bukan merupakan himpunan semua kejadian yang mungkin; dalam hal ini, tidak diperlukan untuk mengetahui atau melakukan observasi terhadap semua kemungkinan
    - Setiap cara yang mungkin terjadi (dalam *n* tersebut) tidak harus memiliki kemungkinan yang sama untuk terjadi.

Probabilitas

- Definisi 2: butuh berapa *n*?
  - Contoh
    - Pada 2 set pengamatan (sampel) yang tidak saling terkait/tergantung, perkiraan probabilitas kejadian *A* dapat ditetapkan berdasarkan masing-masing sampel tersebut.
    - Kedua nilai probabilitas tidak selalu sama satu dengan yang lain.
    - Kedua nilai probabilitas tidak selalu sama dengan perkiraan probabilitas *A* yang ditetapkan dengan pengamatan sejumlah tak-berhingga kali.
  - Problem: berapa jumlah pengamatan *n* yang diperlukan untuk mendapatkan estimasi probabilitas *A* yang dapat diterima?

Probabilitas

- Kisaran (*range*) probabilitas
  - Dari kedua definisi, kisaran probabilitas adalah 0 s.d. 1.
  - prob(A) = 0 "hampir" tidak mungkin terjadi (*nearly impossible*)
  - prob(A) = 1 "hampir" pasti terjadi (*almost certain*)

Probabilitas

- Misal suatu eksperimen (proses) menghasilkan sejumlah output yang berupa variable random
  - Himpunan semua hasil yang mungkin didapat disebut *sample space*.
  - Setiap elemen di dalam sample space disebut *sample points* (*element*)
  - Setiap elemen di dalam sample space memiliki faktor/bobot/*weight* (positif) sedemikian hingga jumlah *weight* seluruh elemen bernilai 1.
  - Nilai bobot berbanding lurus dengan kemungkinan eksperimen akan memberikan hasil elemen tersebut.
  - Bobot tidak lain adalah probabilitas.

Statistika Teknik

Sample Space
Sample Elements

Sample Space dan *Sample Elements*

- Contoh #1
  - Suatu DAS memiliki 3 stasiun: Sta-1, Sta-2, Sta-3.
  - Eksperimen: meneliti setiap stasiun perlu/tidak kalibrasi
  - Output: (y,n,y)
    Sta-1 perlu kalibrasi (y = yes)
    Sta-2 tak perlu kalibrasi (n = no)
    Sta-3 perlu kalibrasi (y = yes)

Sample Space dan *Sample Elements*

- *Sample space*: Alternatif 1
  - S_1={(y,y,y),(y,y,n),(y,n,y),(n,y,y),
    (y,n,n),(n,y,n),(n,n,y),(n,n,n)}
  - S_1 adalah *discrete sample space*: jumlah elemen di dalam S_1 dapat dihitung.
  - Apabila eksperimen dilakukan satu kali saja, maka salah satu elemen S_1 pasti terjadi.

- *Sample space*: Alternatif 2
  - S_2={0,1,2,3}
  - S_2 adalah *discrete sample space*.
  - Hanya ingin diketahui jumlah stasiun yang perlu dikalibrasi.
  - Tidak diperlukan untuk mengetahui stasiun mana yang perlu dikalibrasi.
  - Informasi yang diperoleh lebih sedikit daripada S_1.

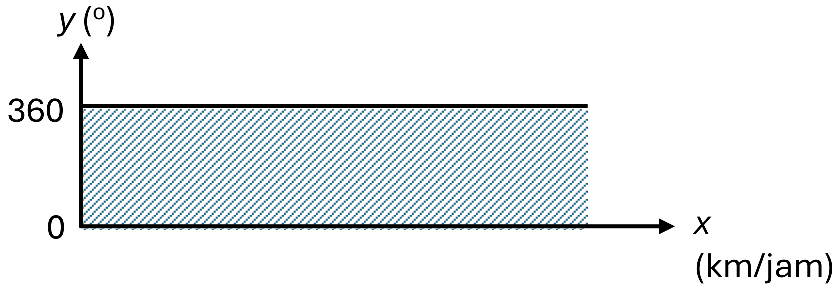
Sample Space dan *Sample Elements*

- Contoh #2
  - Pengukuran angin
    - kecepatan (km/jam) dan
    - arah (°).
  - Output: (x,y)
    - x = kecepatan (km/jam)
    - y = arah (°)

Sample Space dan *Sample Elements*

- *Sample space*: Alternatif 1

$$\Omega_1 = \{(x, y): x \geq 0, 0 \leq y \leq 360\}$$ *continuous sample space*



- Sample space: Alternatif 2

$$\Omega_2 = \{+, -\}$$ *discrete sample space*

  - + = kecepatan > 60 (km/jam)
  - − = kecepatan < 60 (km/jam)

Events

- *Event* adalah suatu himpunan bagian (*subset*) dari *sample space*
- Suatu *event* terjadi jika dan hanya jika hasil dari eksperimen adalah anggota *event* tersebut
- Contoh: Kalibrasi Sta-1, Sta-2, Sta-3
  - *Event A*: paling sedikit 2 stasiun perlu dikalibrasi
    $A=\{(y,y,y),(y,y,n),(y,n,y),(n,y,y)\}$
  - *Event B*: tak ada stasiun yang perlu dikalibrasi
    $B=\{(n,n,n)\}$
  - *Event C*: 2 stasiun perlu dikalibrasi
    $C=\{(y,y,n),(y,n,y),(n,y,y)\}$

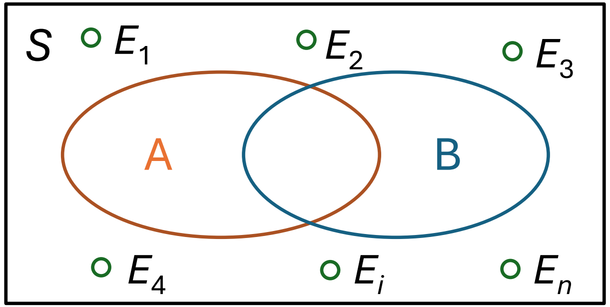
Diagram Venn

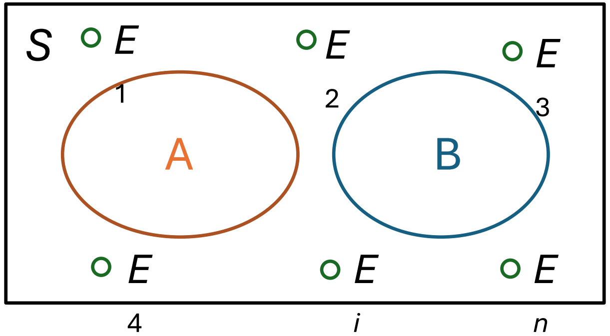
- Notasi:
  S = *sample space*
  E_i = elemen di dalam S
  A,B = *events* di dalam S
  prob(E_i) = probabilitas elemen E_i

$$0 \le \text{prob}(E_1) \le 1$$

$$S = \bigcup_i E_i$$

$$\text{prob}(S) = \sum \text{prob}(E_i) = 1$$





Probabilitas Suatu *Event*

- *Event A*

$$A = \bigcup_{i=m}^{n} E_i$$

$$0 \leq \text{prob}(A) = \sum_{i=m}^{n} \text{prob}(E_i) \leq 1$$

- *Event A* dan *B*

$$\text{prob}(A \cup B) = \text{prob}(A) + \text{prob} - \text{prob}(A \cap B)$$

- Apabila *A* dan *B* tak bergantung satu dengan yang lainnya (*independent*), maka

$$\text{prob}(A \cup B) = \text{prob}(A) + \text{prob}$$

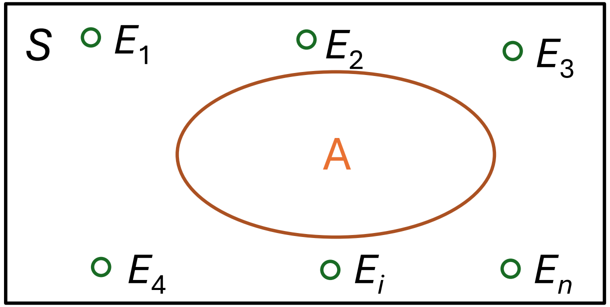
Probabilitas Suatu *Event*

- *Event* A^c (= komplemen *event* A)

$$\text{prob}(A \cap A^c) = 0$$

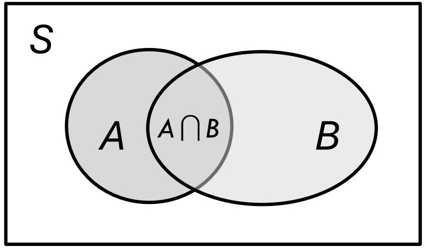
$$\text{prob}(A \cap A^c) = \text{prob}(A) + \text{prob}(A^c) = 1$$

$$\text{prob}(A) = 1 - \text{prob}(A^c)$$



Probabilitas Bersyarat (Conditional Probability)

- Probabilitas suatu event (*event* B) bergantung pada terjadinya *event* lain (*event* A).



prob(B|A) = prob(B) dengan syarat *event* A terjadi

- *sample space* berubah dari S menjadi A,
- *event* diwakili oleh A∩B

$$\mathrm{prob}(B|A) = \frac{\mathrm{prob}(A \cap B)}{\mathrm{prob}(A)}, \qquad \mathrm{prob}(A) \neq 0$$

$$\mathrm{prob}(A \cap B) = \mathrm{prob}(A) \cdot \mathrm{prob}(B|A)$$

Probabilitas Bersyarat (*Conditional Probability*)

- Apabila *event* B tak bergantung pada *event* A (keduanya merupakan *independent events*), maka:
  - prob(B|A) = prob(B), dan
  - prob(A∩B) = prob(A)·prob(B)

Probabilitas Bersyarat (*Conditional Probability*)

- Contoh
  - Data pengamatan hari hujan di suatu wilayah menunjukkan probabilitas hari hujan sbb.
    hari hujan setelah hari hujan = 0.444
    hari tak hujan setelah hari hujan = 0.556
    hari tak hujan setelah hari tak hujan = 0.724
    hari hujan setelah hari tak hujan = 0.276
  - Apabila dijumpai bahwa suatu hari terjadi hujan, berapakah probabilitas bahwa 2 hari berikutnya juga hujan?

Probabilitas Bersyarat (*Conditional Probability*)

- Penyelesaian
  - Misal hari hujan (*h*) terjadi sbb.
    hari ke-0 hari ke-1 hari ke-2
    h *h* *h*

    event *A*

    event *B*
  - *Event A* = hari ke-1 hujan setelah hari ke-0 hujan
    Event B = hari ke-2 hujan setelah hari ke-0 hujan
  - Yang dicari adalah 3 hari hujan berturut-turut:

$$\text{prob}(A \cap B) = \text{prob}(A) \cdot \text{prob}(B|A)$$

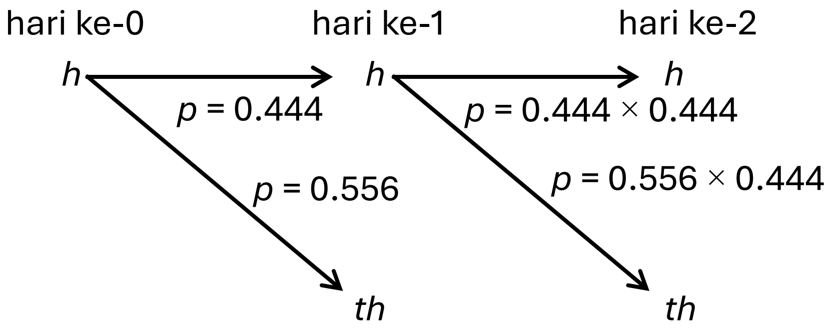
  - Diketahui prob(*A*) = 0.444

$\text{prob}B|A = 0.444$ (hari hujan setelah hari hujan)

$$\text{prob}(A \cap B) = 0.444 \times 0.444 = 0.197$$

Probabilitas Bersyarat (*Conditional Probability*)

- Cara penyelesaian yang lain
  - Probabilitas hari hujan setelah hari hujan adalah $p = 0.444$
  - Suatu hari (hari ke-0) terjadi hujan

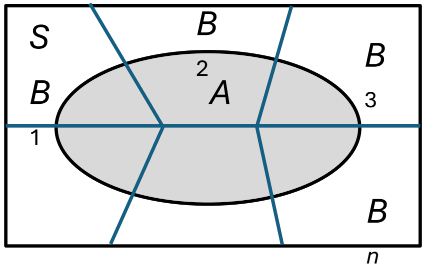


Probabilitas Total (*Total Probability*)

- Apabila B_1, B_2,..., B_n adalah serangkaian *events* yang tidak saling berkaitan (*mutually exclusive events*) dan masing-masing memiliki probabilitas tidak sama dengan nol, prob(B_i) ≠ 0, untuk semua *i*
  - $B_1 \cup B_2 \cup \ldots \cup B_n = S$
  - $B_i \cap B_j = 0, \forall i,j \ (i \neq j)$
  - $\text{prob}(B_i) > 0, \forall i$

Probabilitas Total (*Total Probability*)

- Probabilitas suatu *event A* dapat dituliskan sbb.



$$A = (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_n)$$

$$\text{prob}(A) = \text{prob}[(A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_n)]$$

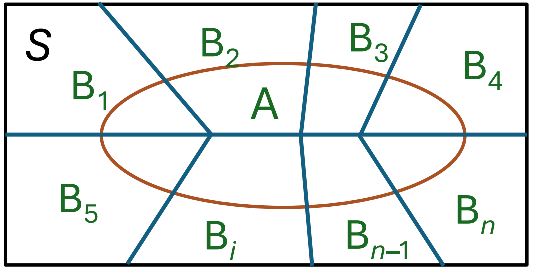
$$\text{prob}(A) = \text{prob}(A \cap B_1) + \text{prob}(A \cap B_2) + \cdots + \text{prob}(A \cap B_n)$$

$$\text{prob}(A) = \text{prob}(B_1 \cap A) + \text{prob}(B_2 \cap A) + \cdots + \text{prob}(B_n \cap A)$$

Probabilitas Total (*Total Probability*)

- Dari probabilitas bersyarat (*conditional probability*)

$$\left.\begin{aligned} \text{prob}(A \cap B_1) &= \text{prob}(A)\text{prob}(B_1|A) \\ \text{prob}(B_1 \cap A) &= \text{prob}(B_1)\text{prob}(A|B_1) \end{aligned}\right\} \quad \begin{aligned} &\text{prob}(A \cap B_1) = \text{prob}(B_1 \cap A) \Longrightarrow \\ &\text{prob}(A)\text{prob}(B_1|A) = \text{prob}(B_1)\text{prob}(A|B_1) \end{aligned}$$



$$\text{prob}(A) = \text{prob}(A \cap B_1) + \text{prob}(A \cap B_2) + \cdots + \text{prob}(A \cap B_n)$$

$$\text{prob}(A) = \text{prob}(B_1)\text{prob}(A|B_1) + \cdots + \text{prob}(B_n)\text{prob}(A|B_n)$$

$$\text{prob}(A) = \sum_{i=1}^{n} \text{prob}(B_i)\text{prob}(A|B_i)$$

Probabilitas Total (*Total Probability*)

- Contoh
  - Data genangan di suatu wilayah permukiman menunjukkan bahwa probabilitas terjadinya genangan adalah 0.80 saat hari hujan dan 0.25 saat tak hujan.
  - Diketahui bahwa probabilitas hari hujan adalah 0.36.
  - Berapakah probabilitas terjadinya genangan di wilayah tersebut?

Probabilitas Total (*Total Probability*)

- Penyelesaian

$$\text{prob}(G) = \text{prob}(H)\text{prob}(G|H) + \text{prob}(H^c)\text{prob}(G|H^c)$$

$$= 0.36\times0.82+(1-0.36)\times0.25$$

$$= 0.448$$

Teorema Bayes

- Dari *conditional probability*

$$\text{prob}(A \cap B) = \text{prob}(A) \cdot \text{prob}(B|A) \qquad \text{.........} (1)$$

$$\text{prob}(B \cap A) = \text{prob}(B) \cdot \text{prob}(A|B) \qquad \text{.........} (2)$$

  - Karena prob(A ∩ B) = prob(B ∩ A), maka

$$\text{prob}(A) \cdot \text{prob}(B|A) = \text{prob}(B) \cdot \text{prob}(A|B) \qquad \text{.........} (3)$$

  - Untuk *event* A dan *event* B_j, persamaan di atas menjadi

$$\text{prob}(A) \cdot \text{prob}(B_j|A) = \text{prob}(B_j) \cdot \text{prob}(A|B_j) \qquad \text{.........} (4)$$

Teorema Bayes

- Dari *total probability*

$$\text{prob}(A) = \sum_{i=1}^{n} \text{prob}(B_i) \cdot \text{prob}(A|B_i) \qquad \text{.........} (5)$$

  - Dengan (5) → (4)

$$\text{prob}\left(B_j|A\right) = \frac{\text{prob}\left(B_j\right) \cdot \text{prob}\left(A|B_j\right)}{\sum_{i=1}^{n} \text{prob}(B_i) \cdot \text{prob}(A|B_i)} \qquad \text{.........} (6)$$

Teorema Bayes

- Pemakaian
  - Untuk mencari probabilitas *event* B_j apabila diketahui *event* A telah terjadi.
  - Untuk mencari (memperkirakan) probabilitas suatu *event* (B_j) dengan mengamati *event* kedua (A).

Teorema Bayes

- Contoh
  - Informasi ramalan cuaca biasa dikirimkan melalui 4 saluran R_i (i = 1,2,3,4) adalah event bahwa informasi tsb dikirimkan melalui saluran i.
  - Probabilitas masing-masing event R_i adalah 0.1, 0.2, 0.3, dan 0.4.
  - Diketahui juga bahwa probabilitas terjadinya kesalahan pengiriman (*event E*) melalui masing-masing saluran adalah: 0.10, 0.15, 0.20, dan 0.25.
  - Suatu saat diketahui bahwa suatu kesalahan pengiriman telah terjadi.
  - Berapakah probabilitas bahwa kesalahan tersebut terjadi melalui saluran ke-2?

Teorema Bayes

- Penyelesaian
  - Diketahui
    prob(R_1) = 0.1 prob($E|R_1$) = 0.10
    prob(R_2) = 0.2 prob($E|R_2$) = 0.15
    prob(R_3) = 0.3 prob($E|R_3$) = 0.20
    prob(R_4) = 0.4 prob($E|R_4$) = 0.25
  - Probabilitas bahwa pengiriman dilakukan melalui saluran ke-2 dengan melihat kenyataan bahwa telah terjadi kesalahan

$$\mathrm{prob}(R_2|E) = \frac{\mathrm{prob}(R_2) \cdot \mathrm{prob}(E|R_2)}{\sum_{i=1}^{n} \mathrm{prob}(R_i) \cdot \mathrm{prob}(E|R_i)}$$

$$\mathrm{prob}(R_2|E) = \frac{0.2 \times 0.15}{0.1 \times 0.10 + 0.2 \times 0.15 + 0.3 \times 0.20 + 0.4 \times 0.25} = 0.15$$

Teorema Bayes

i	prob(R_i)	prob($E \mid R_i$)	prob(R_i) prob($E \mid R_i$)	prob($R_i \mid E$)
1	0.1	0.10	0.01	0.05
2	0.2	0.15	0.03	0.15
3	0.3	0.20	0.06	0.30
4	0.4	0.25	0.10	0.50
∑	**1.0**		**0.20**	**1.00**

prob(E)

Terima Kasih