



twitter: @kh_notodiputro



IPB University
— Bogor Indonesia —

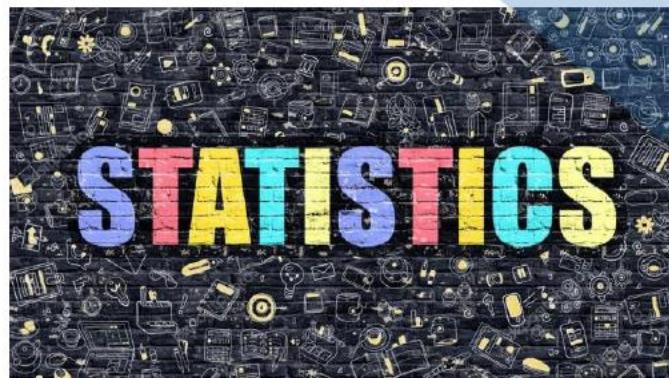
Department of Statistics
Study Program in Statistics and Data Science

Kaidah Peluang dan Peluang Bersyarat

Prof. Dr. Ir. Khairil Anwar Notodiputro, MS

email: khairil@apps.ipb.ac.id

Ketua Program Studi
Statistika dan Sains Data



Kuliah 03

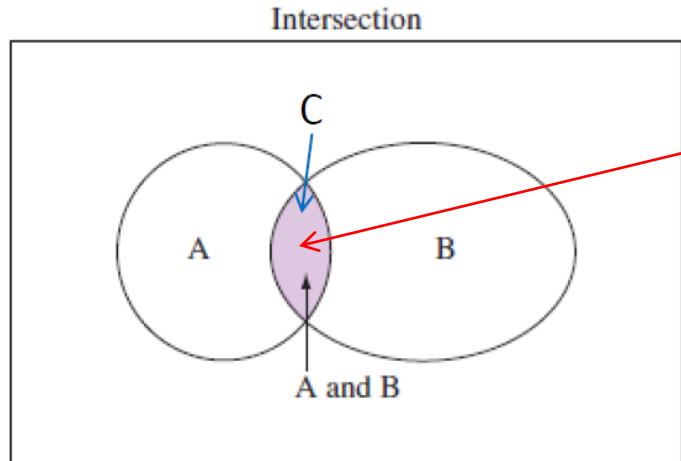


Outline

- **1. Kaidah Peluang**
- **2. Kejadian Tidak Bebas**
- **3. Peluang Bersyarat**
- **4. Dalil Bayes**
- **5. Generalisasi Dalil Bayes**



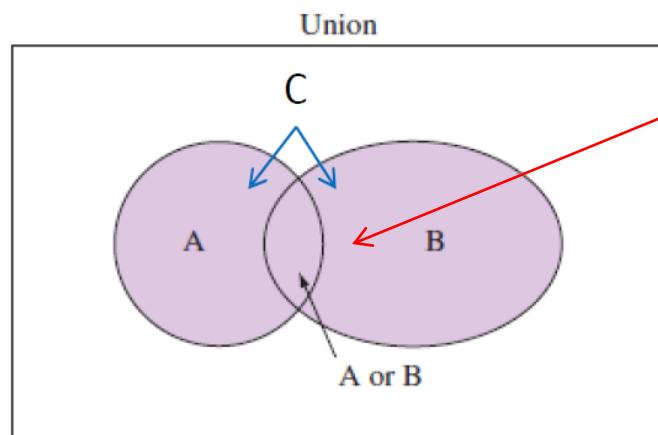
- Irisan dari dua kejadian (*intersection*)



The **intersection** of A and B consists of outcomes that are in both A and B.

- Suatu kejadian C merupakan irisan (*intersection*) dari kejadian A dan B jika dan hanya jika setiap unsur dari C ada di dalam A dan sekaligus ada di dalam B

- Gabungan dari dua kejadian (*union*)



The **union** of A and B consists of outcomes that are in A or B or both.

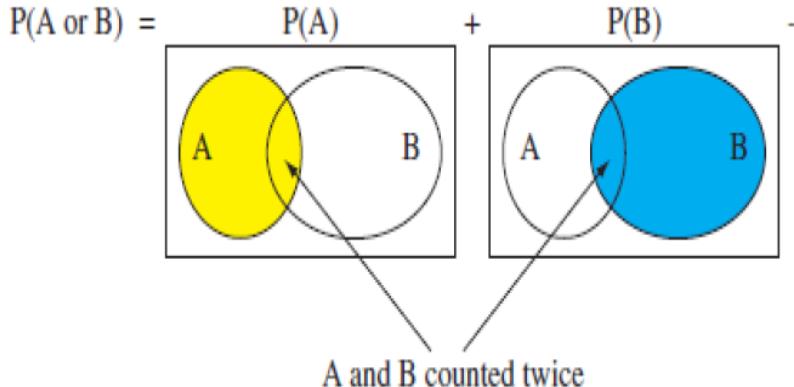
- Suatu kejadian C merupakan gabungan (*union*) dari kejadian A dan B jika dan hanya jika setiap unsur dari C ada di dalam A atau di dalam B atau di dalam A dan sekaligus di dalam B.

Kaidah Peluang

- Peluang gabungan 2 kejadian

Harus dikurangi karena
double counting

For the **union** of two events, $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$.
If the events are **disjoint**, then $P(A \text{ and } B) = 0$, so $P(A \text{ or } B) = P(A) + P(B)$.



- Peluang irisan 2 kejadian

For the **intersection** of two **independent** events, A and B,

$$P(A \text{ and } B) = P(A) \times P(B).$$

Ini hanya berlaku jika kejadian A dan B **bebas**: tejadinya A tidak tergantung B dan sebaliknya.



Ilustrasi



1. Sebuah dadu dilempar, maka **ruang contohnya**:

$$S = \{1, 2, 3, 4, 5, 6\}, n(S)=6$$

jika setiap sisi seimbang maka peluangnya

$$p(1)=p(2)=\dots=p(6)=1/6$$

2. Perhatikan kejadian A yaitu sisi yang muncul kurang atau sama dengan empat maka **ruang kejadiannya**:

$$A = \{1, 2, 3, 4\}, n(A) = 4$$

Maka peluang kejadian A adalah:

$$P(A) = 4/6 = 2/3$$



Ilustrasi Menghitung Peluang



In a telephone survey of 1000 adults, respondents were asked their opinion about the cost of a college education. The respondents were classified according to whether they currently had a child in college and whether they thought the loan burden for most college students is too high, the right amount, or too little. The proportions responding in each category are shown in the **probability table** in Table 4.5. Suppose one respondent is chosen at random from this group.

Probability Table

	Too High (A)	Right Amount (B)	Too Little (C)
Child in College (D)	.35	.08	.01
No Child in College (E)	.25	.20	.11

1. What is the probability that the respondent has a child in college?
2. What is the probability that the respondent does not have a child in college?
3. What is the probability that the respondent has a child in college or thinks that the loan burden is too high or both?



Misal A = kejadian resp merasa biaya PT mahal; D = kejadian resp punya anak di PT

1. $P(D) = 0.35 + 0.08 + 0.01 = 0.44$
2. $P(D^c) = 1 - P(D) = 0.56$
3. $P(A \text{ atau } D) = P(A) + P(D) - P(A \text{ dan } D) = (0.35 + 0.25) + 0.44 - 0.35 = 0.69$

Solution Table 4.5 gives the probabilities for the six simple events in the cells of the table. For example, the entry in the top left corner of the table is the probability that a respondent has a child in college *and* thinks the loan burden is too high ($A \cap D$).

1. The event that a respondent has a child in college will occur regardless of his or her response to the question about loan burden. That is, event D consists of the simple events in the first row:

$$P(D) = .35 + .08 + .01 = .44$$

In general, the probabilities of *marginal* events such as D and A are found by summing the probabilities in the appropriate row or column.

2. The event that the respondent does not have a child in college is the complement of the event D denoted by D^c . The probability of D^c is found as

$$P(D^c) = 1 - P(D)$$

Using the result of part 1, we have

$$P(D^c) = 1 - .44 = .56$$

3. The event of interest is $P(A \cup D)$. Using the Addition Rule

$$\begin{aligned}P(A \cup D) &= P(A) + P(D) - P(A \cap D) \\&= .60 + .44 - .35 \\&= .69\end{aligned}$$

Ilustrasi



An oil-prospecting firm plans to drill two exploratory wells. Past evidence is used to assess the possible outcomes listed in Table .4.

Outcomes for Oil-Drilling Experiment

Event	Description	Probability
A	Neither well produces oil or gas	.80
B	Exactly one well produces oil or gas	.18
C	Both wells produce oil or gas	.02

Dari dua mesin yang beroperasi didefinisikan kejadian ini
A: Tidak satu pun mesin memproduksi minyak atau gas
B: Tepat satu mesin memproduksi minyak atau gas
C: Kedua mesin memproduksi minyak dan gas

Find $P(A \cup B)$ and $P(B \cup C)$.

Solution By their definition, events A, B, and C are jointly mutually exclusive because the occurrence of one event precludes the occurrence of either of the other two. Therefore,

$$P(A \cup B) = P(A) + P(B) = .80 + .18 = .98$$

and

$$P(B \cup C) = P(B) + P(C) = .18 + .02 = .20$$

$$\begin{aligned} A \cap B &= \emptyset \\ A \cap C &= \emptyset \\ B \cap C &= \emptyset \end{aligned}$$



Ilustrasi

Dalam satu keranjang terdapat 10 jeruk yang terdiri 4 jeruk warna hijau dan 6 jeruk warna kuning. Santi memilih secara acak 3 buah jeruk. Berapa peluang:

- a) jeruk yang terpilih semuanya berwarna kuning?
- b) Dua jeruk kuning dan satu jeruk hijau

Solusi:

- a) Misalkan kejadian A = jeruk yang terpilih semuanya berwarna kuning

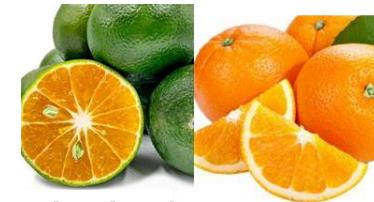
✓ Anggota ruang contoh:

$$n(S) = \binom{10}{3} = \frac{10!}{3! \times 7!} = \frac{10 \times 9 \times 8}{3 \times 2 \times 1} = 120$$

✓ Anggota ruang kejadian:

$$n(A) = \binom{6}{3} = \frac{6!}{3! \times 3!} = \frac{6 \times 5 \times 4}{3 \times 2 \times 1} = 20$$

✓ Maka $P(A) = \frac{20}{120} = \frac{1}{6}$



- b) Misalkan kejadian B = Dua jeruk kuning dan satu jeruk hijau

✓ Anggota ruang contoh:

$$n(S) = \binom{10}{3} = \frac{10!}{3! \times 7!} = \frac{10 \times 9 \times 8}{3 \times 2 \times 1} = 120$$

✓ Anggota ruang kejadian:

$$\begin{aligned} n(B) &= \binom{6}{2} \binom{4}{1} = \left(\frac{6!}{2! \times 4!} \right) \left(\frac{4!}{1! \times 3!} \right) \\ &= \left(\frac{6 \times 5}{2 \times 1} \right) \left(\frac{4}{1} \right) = 60 \end{aligned}$$

✓ Maka $P(B) = \frac{60}{120} = \frac{1}{2}$



Diskusi Dulu.....

Sesi 1... beres!!!

Sesi 2...



Kejadian tidak bebas



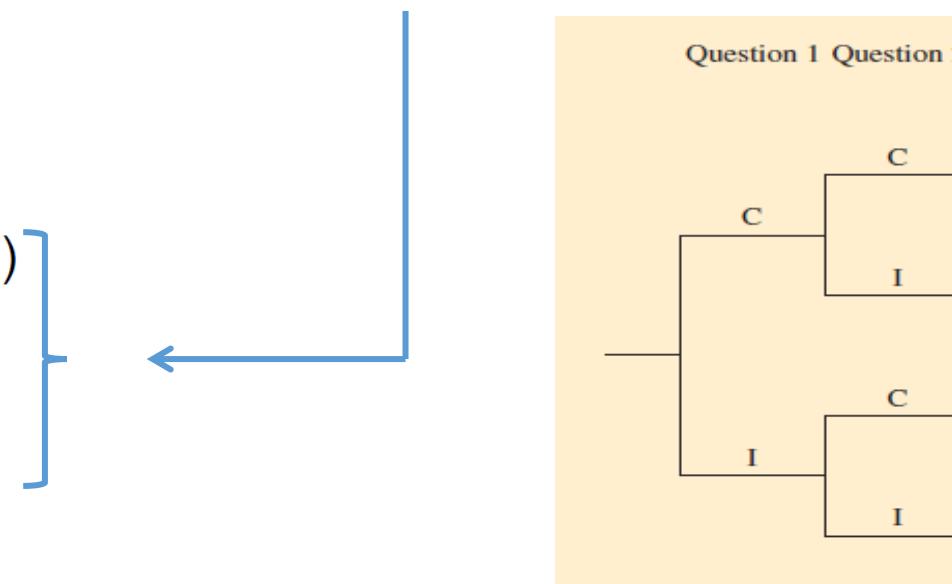
- Bagaimana jika 2 kejadian tdk bebas?

2nd Question		
1st Question	C	I
C	0.58	0.05
I	0.11	0.26
A and B		

- $P(A) = P(\{CC, CI\}) = 0.58 + 0.05 = 0.63$
- $P(B) = P(\{CC, IC\}) = 0.58 + 0.11 = 0.69$
- $P(AB) = P(\{CC\}) = 0.58$ (ini dilihat dalam tabel)
- $P(A) \times P(B) = 0.63 \times 0.69 = 0.43$
- A dan B tdk bebas $\leftrightarrow P(AB) \neq P(A) \times P(B)$

Perhatikan apakah kejadian seorang mahasiswa menjawab soal nomor 1 dan nomor 2 benar merupakan dua kejadian yang bebas?

- Misal A = soal nomor 1 dijawab dengan benar, B= soal nomor 2 dijawab dengan benar.
- Jika kejadian A dan B bebas maka harus berlaku bahwa $P(AB) = P(A) \times P(B)$.
- Mari kita periksa apakah hal itu berlaku?



Kejadian tidak bebas



- Perhatikan uji diagnostik untuk mengetahui pengguna narkoba:
- Jika diketahui bhw hasil ujinya positif, berapa peluang bhw ybs benar pengguna narkoba?

$P(\text{narkoba jika uji positif})$

- Peluang A **bersyarat** B mrp peluang terjadinya A jika diketahui B telah terjadi → ingat B terjadi lebih dulu daripada A.
- Besarnya adalah

$$P(A|B) = \frac{P(AB)}{P(B)}$$

Berlaku juga jika dibalik

$$P(B|A) = \frac{P(AB)}{P(A)}$$

Dibaca “**peluang B bersyarat A**”.

- Ini dikenal sebagai peluang bersyarat, atau dibaca sbg “**peluang A bersyarat B**”.



Peluang Bersyarat

CONDITIONAL PROBABILITIES

The conditional probability of event A , given that event B has occurred is

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

The conditional probability of event B , given that event A has occurred is

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

Notice that, in this form, you need to know $P(A \cap B)$!

Dari pengamatan bertahun-tahun terhadap masyarakat di suatu daerah diketahui ada 51% laki-laki dan 49% perempuan. Proporsi orang yang menderita buta warna adalah sbb:

	Men(B)	Women (B^C)	Total
Colorblind (A)	.04	.002	.042
Not Colorblind (A^C)	.47	.488	.958
Total	.51	.49	1.00

Jika diambil secara acak 1 orang dan ternyata laki-laki (B), berapa peluang orang ini menderita buta warna (A)?



Ini masalah peluang bersyarat

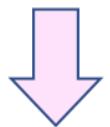
Yang dipertanyakan adalah berapa $P(A|B)$.

$$\begin{aligned} P(A|B) &= \frac{P(A \text{ dan } B)}{P(B)} \\ &= \frac{0.04}{0.51} \quad (\text{lihat tabel}) \\ &= 0.078 \end{aligned}$$



Ilustrasi

A study² of 5282 women aged 35 or over analyzed the Triple Blood Test to test its accuracy. It was reported that of the 5282 women, “48 of the 54 cases of Down syndrome would have been identified using the test and 25 percent of the unaffected pregnancies would have been identified as being at high risk for Down syndrome (these are false positives).”



Tabel Kontingensi

Down Syndrome Status	Blood Test		
	POS	NEG	Total
D (Down)	48	6	54
D ^c (unaffected)	1307	3921	5228
Total	1355	3927	5282

Department of Statistics
Statistics and Data Science Study Program

- $P(D|POS)$: peluang benar *down syndrome* jika hasil uji positif
 - $P(D|POS) = 48/1355 = 0.035 \leftarrow$ langsung dari tabel
 - $P(D|POS) = P(D \text{ dan } POS)/P(POS) \leftarrow$ rumus
 - $P(D \text{ dan } POS) = 48/5282 = 0.0091 \leftarrow$ lihat tabel
 - $P(POS) = 1355/5282 = 0.257 \leftrightarrow P(D|POS) = 0.035$
- Ternyata jika hasil uji seorang wanita positif maka peluang dia benar terkena *syndrome* <4%.
- Ini contoh uji yang buruk.

Peluang Bersama → P(A dan B)



- Sebelumnya jika A dan B bebas:
 - $P(A \text{ dan } B) = P(A) P(B)$
 - Dengan peluang bersyarat terbuka hal yg lbh umum
- Untuk kejadian A dan B (bebas atau tdk bebas):
 - $P(A \text{ dan } B) = P(A) P(B|A)$
 - Berlaku juga $P(A \text{ dan } B) = P(B) P(A|B)$
 - Jika A dan B bebas maka $P(B|A) = P(B)$, begitu pula $P(A|B) = P(A)$
- Di dalam kotak ada 4 bola hitam dan 6 merah:
 - H_i = pengambilan ke i dapat hitam, M_j = pengambilan ke j dapat merah
 - Brp $P(H_1)$? Berapa $P(M_2|H_1)$? Berapa $P(H_1M_2)$?



Pemulihan vs Tanpa Pemulihan



- Di dalam kotak ada 4 bola hitam dan 6 merah:
 - H_i = pengambilan ke i dapat hitam, M_j = pengambilan ke j dapat merah
 - Brp $P(H_1)$? Berapa $P(M_2|H_1)$? Berapa $P(H_1M_2)$?
- Dua skenario:
 - Tanpa pemulihan (*without replacement*) yaitu bola yang terambil tdk dikembalikan lagi ke dalam kotak, shg $P(H_1) = 4/10$, $P(M_2|H_1) = 6/9$, dan $P(H_1M_2) = P(H_1) P(M_2|H_1) = (4/10)(6/9) = 24/90$
 - Dengan pemulihan (*with replacement*) yaitu bola yang terambil dikembalikan lagi ke dalam kotak, shg $P(H_1) = 4/10$, $P(M_2|H_1) = 6/10$, dan $P(H_1M_2) = P(H_1) P(M_2|H_1) = (4/10)(6/10) = 24/100$



Peluang	T. Pemulihan	D. Pemulihan
H_1	$4/10$	$4/10$
$M_2 H_1$	$6/9$	$6/10$
H_1M_2	$24/90$	$24/100$



Memeriksa Kebebasan 2 Kejadian



Status	Blood Test		Total
	POS	NEG	
D	0.009	0.001	0.010
D^c	0.247	0.742	0.990
Total	0.257	0.743	1.00

- Apakah D dan POS bebas? $P(POS|D)$ vs $P(POS)$
 - $P(POS|D) = P(POS \text{ dan } D) / P(D) = 0.009/0.010 = 0.9$
 - $P(POS) = 0.257 \rightarrow \text{jadi D dan POS tidak bebas!!!}$
- Apakah D^c dan POS bebas? \rightarrow Pikirkan!

Karena $P(POS|D) \neq P(POS)$,
artinya terjadinya POS
tergantung pada terjadinya D.

Here are three ways to determine if events A and B are independent:

- Is $P(A|B) = P(A)$?
- Is $P(B|A) = P(B)$?
- Is $P(A \text{ and } B) = P(A) \times P(B)$?

If any of these is true, then the others are also true and the events A and B are independent.



Diskusi Dulu.....

Sesi 2... beres!!!

Sesi 3...



Dalil Bayes



- Dalam praktik sering ingin diketahui berapa peluang benar syndrome jika hasil uji positif? → $P(D|POS) = ?$
- Untuk ini perlu info brp $P(POS)$ dan $P(D)$
- Thomas Bayes (1763) membuat dalil:
 - Jika A dan B adalah dua kejadian dengan peluang tdk sama dengan 0 atau 1, maka

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

- Jadi $P(B|A)$ dpt dicari tanpa harus tahu $P(AB)$ dan $P(A)$.
- Sehingga untuk pertanyaan di atas, yaitu $P(D|POS)$ dapat ditulis:

$$\bullet \quad P(D|POS) = \frac{P(POS|D)P(D)}{P(POS|D)P(D) + P(POS|D^c)P(D^c)}$$

Biasanya $P(B|A) = \frac{P(AB)}{P(A)}$



Ilustrasi



Down Syndrome Status	Blood Test		
	POS	NEG	Total
D (Down)	48	6	54
D ^c (unaffected)	1307	3921	5228
Total	1355	3927	5282

$$P(D|Pos) = \frac{P(Pos|D)P(D)}{P(Pos|D)P(D) + P(Pos|D^c)P(D^c)}$$

$$P(D|Pos) = \frac{(48/54)(54/5282)}{(48/54)(54/5282) + (1307/5228)(5228/5282)} = \frac{0.009087}{(0.009087 + 0.247444)}$$

- Kalau dihitung lebih lanjut maka $P(D|POS) = 0.035$
- Hasilnya persis sama dengan yang sebelumnya.

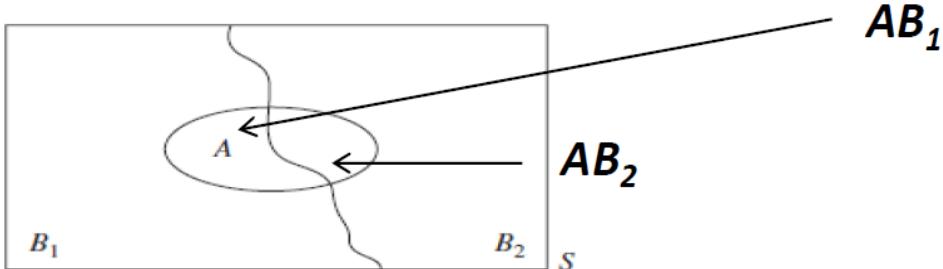


Pembuktian Dalil Bayes



IPB University
Bogor Indonesia

Department of Statistics
Statistics and Data Science Study Program



$$A = AB_1 \cup AB_2.$$

$$P(A) = P(AB_1) + P(AB_2).$$

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2).$$

- Ingat: $P(B_1|A) = \frac{P(AB_1)}{P(A)}$

- Sehingga Dalil Bayes: $P(B_1|A) = \frac{P(A|B_1)P(B_1)}{P(A|B_1)P(B_1) + P(A|B_2)P(B_2)}$

atau

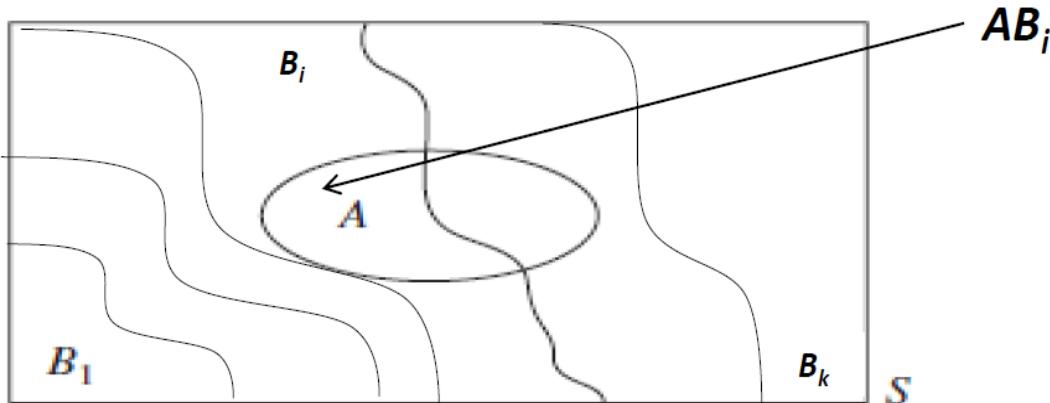
$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$



Thomas Bayes, 1701-1761



Generalisasi Dalil Bayes



AB_i

$$A = AB_1 \cup AB_2 \cup \dots \cup AB_k$$

$$P(A) = P(AB_1) + P(AB_2) + \dots + P(AB_k)$$

$$P(A) = P(B_1)P(A|B_1) + \dots + P(B_k)P(A|B_k)$$

- Secara umum:
$$P(A) = \sum_{i=1}^k P(B_i)P(A|B_i).$$
- Ingat:
$$P(B_i|A) = \frac{P(AB_i)}{P(A)}$$
- Sehingga Dalil Bayes menjadi:

Bayes' Rule. If the events B_1, B_2, \dots, B_k form a partition of the sample space S , and A is any event in S , then

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^k P(A|B_j)P(B_j)}.$$



Ilustrasi Dalil Bayes

Kota Bogor disebut kota hujan karena peluang terjadinya hujan (H) cukup besar yaitu sebesar 0.6. Hal ini menyebabkan para mahasiswa harus siap-siap dengan membawa payung (P). Peluang seorang mahasiswa membawa payung jika hari hujan 0.8, sedangkan jika tidak hujan 0.4.

Berapa peluang hari akan hujan jika diketahui mahasiswa membawa payung?



Misalkan :

H = Bogor hujan,

P = mahasiswa membawa payung

$$P(H) = 0.6 \quad P(TH) = 1 - 0.6 = 0.4 \quad P(P|H) = 0.8$$

$$P(P|TH) = 0.4$$

Ditanya : $P(H|P)$???

Jawab :

Sesuai hukum perkalian peluang



$$P(H / P) = \frac{P(H \cap P)}{P(P)} = \frac{P(H \cap P)}{P(H \cap P) + P(TH \cap P)} = \frac{P(H)P(P / H)}{P(H)P(P / H) + P(TH)P(P / TH)}$$

$$P(H / P) = \frac{0.6 \times 0.8}{0.6 \times 0.8 + 0.4 \times 0.4} = \frac{0.48}{0.48 + 0.16} = \frac{0.48}{0.64}$$

→ Dalil Bayes

Dari Buku Mendenhall

4.6 Preschool or Not? On the first day of kindergarten, the teacher randomly selects 1 of his 25 students and records the student's gender, as well as whether or not that student had gone to preschool.

- How would you describe the experiment?
- Construct a tree diagram for this experiment. How many simple events are there?
- The table below shows the distribution of the 25 students according to gender and preschool experience. Use the table to assign probabilities to the simple events in part b.

	Male	Female
Preschool	8	9
No Preschool	6	2

Simple events dari kejadian A adalah semua unsur dari A

4.51 Suppose that $P(A) = .4$ and $P(A \cap B) = .12$.

- Find $P(B|A)$.
- Are events A and B mutually exclusive?
- If $P(B) = .3$, are events A and B independent?

4.54 Drug Testing Many companies are now testing prospective employees for drug use. However, opponents claim that this procedure is unfair because the tests themselves are not 100% reliable. Suppose a company uses a test that is 98% accurate—that is, it correctly identifies a person as a drug user or nonuser with probability .98—and to reduce the chance of error, each job applicant is required to take two tests. If the outcomes of the two tests on the same person are independent events, what are the probabilities of these events?

- A nonuser fails both tests.
- A drug user is detected (i.e., he or she fails at least one test).
- A drug user passes both tests.

Diskusikan

5.57

TRY

Mammogram diagnostics Breast cancer is the most common form of cancer in women, affecting about 10% of women at some time in their lives. There is about a 1% chance of having breast cancer at a given time (that is, $P(S) = 0.01$ for the state of having breast cancer at a given time). The chance of breast cancer increases as a woman ages, and the American Cancer Society recommends an annual mammogram after age 40 to test for its presence. Of the women who undergo mammograms at any given time, about 1% are typically estimated to actually have breast cancer. The likelihood of a false test result varies according to the breast density and the radiologist's level of experience. For use of the mammogram to detect breast cancer, typical values reported are sensitivity = 0.86 and specificity = 0.88.

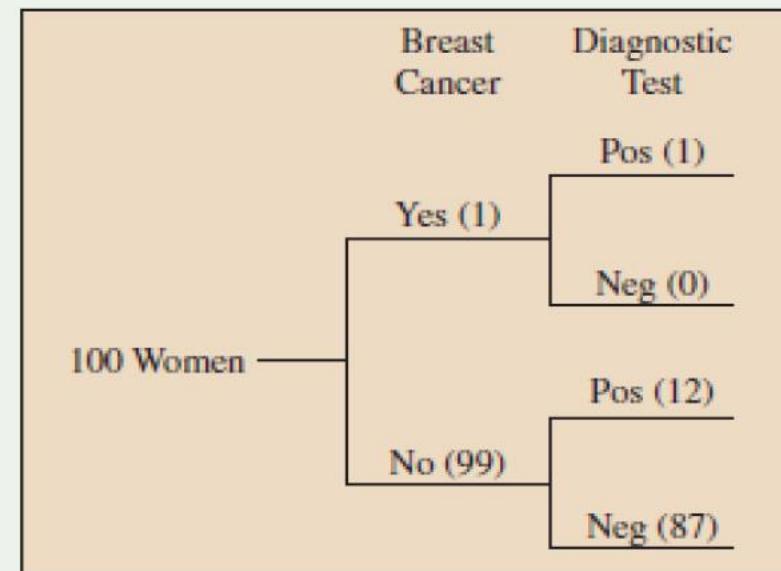
- Construct a tree diagram in which the first set of branches shows whether a woman has breast cancer and the second set of branches shows the mammogram result. At the end of the final set of branches, show that $P(S \text{ and } \text{POS}) = 0.01 \times 0.86 = 0.0086$ and report the other intersection probabilities also.
- Restricting your attention to the two paths that have a positive test result, show that $P(\text{POS}) = 0.1274$.



IPB University
Bogor Indonesia

Department of Statistics
Statistics and Data Science Study Program

- Of the women who receive a positive mammogram result, what proportion actually have breast cancer?
- The following tree diagram illustrates how $P(S | \text{POS})$ can be so small, using a typical group of 100 women who have a mammogram. Explain how to get the frequencies shown on the branches and why this suggests that $P(S | \text{POS})$ is only about 0.08.



Typical results of mammograms for 100 women

Diskusikan



5.116 Bayes' rule Suppose we know $P(A)$, $P(B|A)$, and $P(B^c|A^c)$, but we want to find $P(A|B)$.

- Using the definition of conditional probability for $P(A|B)$ and for $P(B|A)$, explain why $P(A|B) = P(A \text{ and } B)/P(B) = [P(A)P(B|A)]/P(B)$.
- Splitting the event that B occurs into two parts, according to whether A occurs, explain why

$$P(B) = P(B \text{ and } A) + P(B \text{ and } A^c).$$

- Using part b and the definition of conditional probability, explain why

$$P(B) = P(A)P(B|A) + P(A^c)P(B|A^c).$$

- Combining what you have shown in parts a–c, reason that

$$P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A) + P(A^c)P(B|A^c)}.$$

This formula is called **Bayes' rule**. It is named after a British clergyman who discovered the formula in 1763.

4.79 Screening Tests Suppose that a certain disease is present in 10% of the population, and that there is a screening test designed to detect this disease if present. The test does not always work perfectly. Sometimes the test is negative when the disease is present, and sometimes it is positive when the disease is absent. The table below shows the proportion of times that the test produces various results.

	Test Is Positive (P)	Test Is Negative (N)
Disease Present (D)	.08	.22
Disease Absent (D^c)	.05	.85

- Find the following probabilities from the table: $P(D)$, $P(D^c)$, $P(N|D^c)$, $P(N|D)$.
- Use Bayes' Rule and the results of part a to find $P(D|N)$.
- Use the definition of conditional probability to find $P(D|N)$. (Your answer should be the same as the answer to part b.)
- Find the probability of a false positive, that the test is positive, given that the person is disease-free.
- Find the probability of a false negative, that the test is negative, given that the person has the disease.
- Are either of the probabilities in parts d or e large enough that you would be concerned about the reliability of this screening method? Explain.



😊 THANK YOU 😊





Thank You

email: khairil@apps.ipb.ac.id

twitter: @kh_notodiputro

