Program Sarjana S1
Porgram Studi Statistika
dan Sains Data

IPB University
— Bogor Indonesia —

RESPONSI
PENGANTAR STATISTIKA BAYES

Pertemuan 1
Review Peluang Bersyarat dan Rumus Bayes

Kamis, 25 Januari 2024

Pendahuluan : Peluang Bersyarat

Jika diberikan dua kejadian, A dan B, **Peluang bersyarat** A jika kejadian B **telah terjadi** yang disimbolkan P(A|B) dapat didefinisikan:

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

dengan $P(B) > 0$

Jika A dan B Saling bebas maka

$$P(A \cap B) = P(A)P(B)$$

Ilustrasi:
A : Peluang menonton film "*Descendent of the sun*" = 2/3
B : Peluang menonton film "*Ant Man*" = 1/2
Peluang menonton beririsan keduanya 1/3

$$P(A) = 2/3, P(B) = 1/2, P(A \cap B) = 1/3$$

$$P(A|B) = \frac{1/3}{1/2} = \frac{2}{3}$$

$$P(B|A) = \frac{1/3}{2/3} = \frac{1}{2}$$

bagaimana jika saling bebas?

Pendahuluan : Peluang Bersyarat & Rumus Bayes

Jika A saling bebas dengan B, maka

$$P(A \cap B) = P(A)P(B)$$

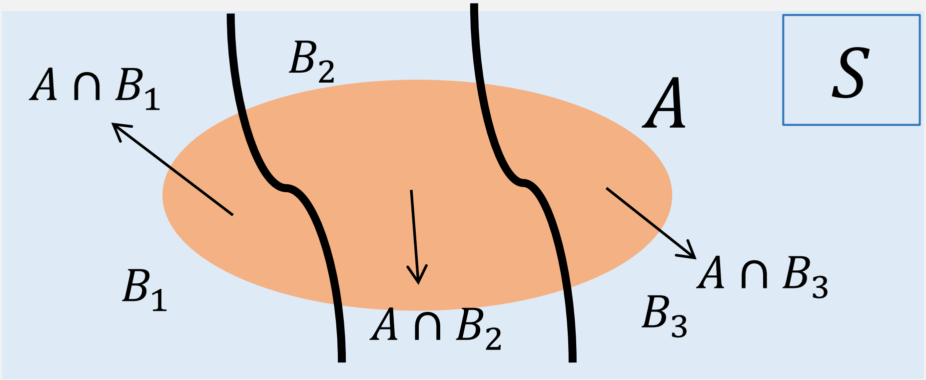
$$P(A|B) = \frac{P(A)P(B)}{P(B)} = P(A)$$

Dari kedua persamaan di atas diperoleh

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

Maka dari definisi peluang bersyarat, diperoleh **rumus bayes** dapat ditulis

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



- Ruang S dipartisi menjadi $B_1, B_2, \ldots, B_k$ dengan
$$B_i \cap B_j = \emptyset$$
- Karena A subset dari S maka
$$A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots \cup (A \cap B_k)$$
$$P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \ldots + P(A|B_k)P(B_k)$$
$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i)$$
- Maka **Rumus bayes** dari peluang bersyarat dapat ditulis
$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}$$

Pendahuluan : Peluang Bersyarat & Rumus Bayes

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

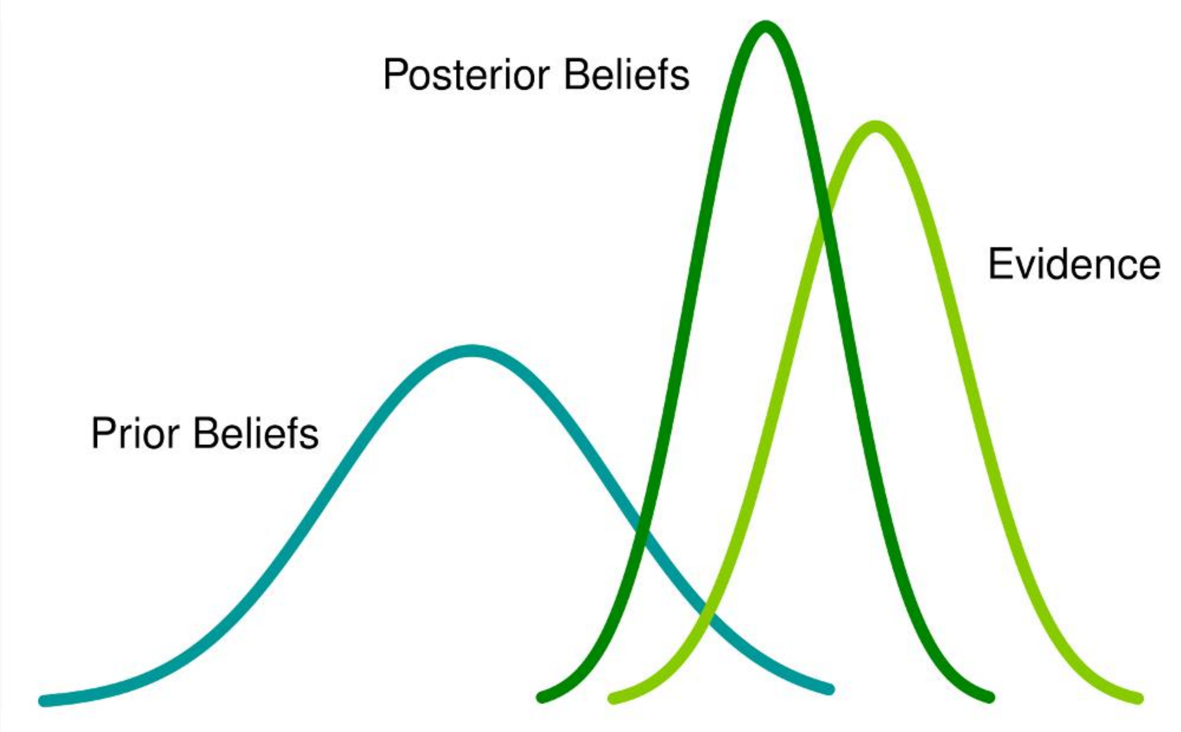
$P(B)$ adalah bukti/evidence (data)

$P(A)$ adalah **prior** sebagai hipotesis

$P(A|B)$ adalah **posterior**

$P(B|A)$ adalah **likelihood**

$$P(hipo|data) = \frac{P(data|hipo)P(hipo)}{P(data)}$$



Pendahuluan : Ilustrasi

Ilustrasi 1

Seorang Dosen Statistika ingin memilih peserta Festival Satria Data 2024 secara acak di dua kelas yang berbeda yakni Kelas Pemodelan(KP) dan Kelas Teori (KT). Setelah itu mengambil secara acak siswa di dalam kelas tersebut. Kondisinya:

Kelas Pemodelan terdapat 12 mahasiswa perempuan dan 18 mahasiswa laki-laki.

Kelas Teori terdapat 20 mahasiswa perempuan dan 15 mahasiswa laki-laki. Tentukan

- Peluang bahwa siswa yang terpilih adalah perempuan
- Peluang bersyarat bahwa Kelas Pemodelan yang terpilih jika siswanya (*given*) adalah perempuan

Pendahuluan : Ilustrasi

Solusi Ilustrasi 1

1. Peluang bahwa siswa yang terpilih adalah perempuan (F).
   - Peluang memilih KP dan KT

$$P(KP) = P(KT) = \frac{1}{2}$$

   - Peluang perempuan terpilih pada KP dan KT

$$P(F|KP) = \frac{12}{30}$$

$$P(F|KT) = \frac{20}{35}$$

$$P(F) = P(F|KP)P(KP) + P(F|KT)P(KT)$$
$$= \frac{12}{30} \times \frac{1}{2} + \frac{20}{35} \times \frac{1}{2} = \frac{17}{35}$$

2. $P(KP|F)$ menggunakan Aturan Bayes

$$P(KP|F) = \frac{P(F|KP)P(KP)}{P(F)}$$

$$P(KP|F) = \frac{\frac{12}{30} \times \frac{1}{2}}{\frac{17}{35}}$$

$$= \frac{12/60}{17/35} = \frac{7}{17}$$

Pendahuluan : Ilustrasi

Ilustrasi 2

Setelah melakukan perampokan, Seorang pencuri melompat ke dalam taksi dan menghilang. Seorang saksi mata di TKP memberi tahu polisi bahwa taksi itu berwarna kuning. Untuk memastikan bahwa kesaksian ini bernilai, Seorang Asisten melakukan analisis Bayesian terhadap situasi tersebut. Setelah melakukan riset lapangan dia mendapatkan informasi berikut:

- Di kota tersebut, 80% taksi berwarna hitam dan 20% taksi berwarna kuning.
- Saksi mata tidak selalu dapat diandalkan dan dari pengalaman masa lalu, diharapkan bahwa seorang saksi mata memiliki keakuratan sebesar 70%. Dia akan mengidentifikasi warna taksi secara akurat (kuning atau hitam) 7 dari 10 kali.

Pendahuluan : Ilustrasi

Solusi Ilustrasi 2

Pertama Asisten melakukan identifikasi dari informasi yang ada. Misalkan

True=Yellow jika warna taxi adalah kuning

True=Black jika warna taxi adalah Hitam

report=Yellow jika warna taxi adalah kuning

report=black jika warna taxi adalah hitam

Tujuan:

$$P(true = Y|report = Y)$$

Dengan menggunakan aturan bayes

$$P(true = Y|report = Y)$$

$$= \frac{P(report = Y|true = Y)P(true = Y)}{P(report = Y)}$$

dari informasi yang ada

$$P(true = Y) = 0{,}2$$

$$P(report = Y|true = Y) = 0{,}7$$

sehingga $P(report = Y)$ dapat ditentukan

$$P(report = Y)$$

$$= P(report = Y|true = Y)P(true = Y)$$

$$+ P(report = Y|true = B)P(true = B)$$

$$= 0{,}7 \times 0{,}2 + 0{,}3 \times 0{,}8 = 0{,}14 + 0{,}24 = 0{,}38$$

Sehingga dipeoleh

$$P(true = Y|report = Y) = \frac{0{,}7 \times 0{,}2}{0{,}38} = 0{,}368$$

sehingga kesaksian para saksi mata tidak memberikan banyak kepastian yang bermanfaat.

Pendahuluan : Ilustrasi (Bayes Box)

Ilustrasi 3

Ada sebuah wadah yang berisi 9 bola. Bola-bola tersebut dapat berwarna **hijau** atau **merah**. Banyaknya bola merah di dalam kotak tidak diketahui. Misalkan X adalah banyaknya bola merah di dalam kotak. Misalkan kita mengamati dua kali pengambilan dari kotak (tanpa pengembalian).

Hasil pengambilan pertama berwarna merah, hasil pengambilan kedua berwarna hijau.

Tentukan distribusi posterior dari X dengan mengisi tabel bayes box berikut.

X	prior	likelihood	prior x likelihood	Posterior

Pendahuluan : Ilustrasi

Solusi Ilustrasi 3

Diketahui:

- X adalah banyaknya bola merah di dalam kotak
- Likelihood: kita mengamati dua kali pengambilan dari kotak (tanpa pengembalian), pengambilan pertama berwarna merah, hasil pengambilan kedua berwarna hijau

Ditanya: distribusi posterior dari X dengan mengisi tabel bayes box

- Berdasarkan table bayes box kita perlu mendefinisikan prior dan likelihood untuk menghitung posterior
- Misal M adalah pengambilan pertama berwarna merah
- Misal H adalah pengambilan kedua berwarna hijau
- Likelihood tersebut dapat dihitung dengan

$$P(M,H|X=x)=\left(\frac{x}{9}\right)\left(\frac{9-x}{8}\right), x=0,1,2,3,\ldots,9$$

Pendahuluan : Ilustrasi

Solusi Ilustrasi 3

- Setelah likelihood, kita perlu mendefinisikan prior dari X. Ingat bahwa prior bersifat subjektif
- Agar mudah misal kita mendefinisikan prior yang sama untuk setiap $X = x$, sedemikian sehingga

$$P(X = x) = \frac{1}{9}, x = 0,1,2,3,\dots,9$$

Pendahuluan : Ilustrasi

Solusi Ilustrasi 3

X	prior	likelihood	priorxlikelihood	posterior
0	1/9	0	0	0
1	1/9	1/9	1/81	1/15
2	1/9	7/36	7/324	7/60
3	1/9	1/4	1/36	3/20
4	1/9	5/18	5/162	1/6
5	1/9	5/18	5/162	1/6
6	1/9	1/4	1/36	3/20
7	1/9	7/36	7/324	7/60
8	1/9	1/9	1/81	1/15
9	1/9	0	0	0
Total			5/27	

Pendahuluan : Question

Tugas 1

Peluang Dena akan menyukai sebuah restoran adalah 0.7. Di antara restoran yang ia sukai, 20% memiliki bintang 5 di aplikasi OKfood, 50% memiliki bintang 4, dan 30% memiliki kurang dari bintang 4. Informasi apa lagi yang kita perlukan jika kita ingin mencari peluang bahwa Dena menyukai suatu restoran jika restoran tersebut memiliki kurang dari bintang 4 di aplikasi OKfood? Tunjukkan Secara Matematis mengapa informasi tersebut diperlukan?

Pendahuluan : Question

Tugas 2

Misalkan Y adalah banyakanya cacahan (*count*) dari peubah acak Poisson (μ).

Misalkan 5 kemungkinan nilai μ adalah 1, 2, 3, 4, dan 5. Kemudian misalkan saja kita tidak memiliki alasan untuk memberikan bobot yang mungkin lebih besar satu nilai dibandingkan dengan nilai yang lain, sehingga kita memberikan bobot yang sama untuk setiap nilai.

Selanjutnya diketahui bahwa $Y = 2$ telah teramati.

a. Tentukan distribusi posterior dengan mengisi tabel bayes box

b. Buatlah soal a dengan menggunakan R

c. Buatlah grafik barchart berdasarkan hasil b dengan sumbu x adalah nilai-nilai μ dan sumbu y adalah nilai peluang posterior

Pendahuluan : Question

Tugas 3

Misalkan n adalah banyaknya pelanggan yang mengunjungi toko pada hari diskon dan banyaknya pelanggan tersebut tidak diketahui. Jumlah pelanggan yang melakukan pembelian adalah $Y|n$~Binomial(n, θ) di mana θ adalah peluang untuk melakukan pembelian jika pelanggan mengunjungi toko tersebut. Nilai priornya adalah n~Poisson(5). Dengan mengasumsikan bahwa θ diketahui dan n adalah parameter yang tidak diketahui, Buat grafik dengan bantuan R distribusi posterior dari n untuk semua kombinasi $Y \in \{0,5,10\}$ dan $\theta \in \{0.2,0.5\}$ dan berikan komentar mengenai pengaruh Y dan posteriornya.


RAHMET
OBRIGADO
GRAZIE
HVALA
谢谢
KIITOS
DANKE
TAK
DANK JE
THANK YOU
SHUKRAAN
ευχαριστώ
GRACIAS
BARKA
MERCI
СПАСИБО
ARIGATO
TAKK
TACK