IPB University
Inspiring Innovation with Integrity

Pengantar Statistika Bayes - STA1312

Konsep Inferensi Bayesian – Part 1

(Bayes' Rule dan Bayes' Box)

Dr. Kusman Sadik, S.Si, M.Si
Program Studi Statistika dan Sains Data IPB
Tahun Akademik 2023/2024

Teorema Bayes

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \begin{cases} P(A \cap B) = P(B) \times P(A|B) \\ P(A \cap \tilde{B}) = P(\tilde{B}) \times P(A|\tilde{B}) \\ P(A) = P(A \cap B) + P(A \cap \tilde{B}) \end{cases}$$

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

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

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

Fakta hubungan ini dapat digunakan untuk menghitung P(B|A) berdasarkan P(A|B)

Teorema Bayes (*Cont.*)

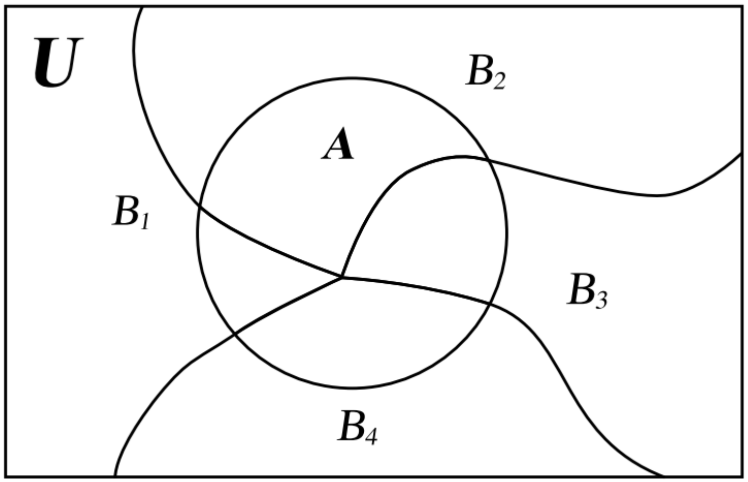
A set of events partitioning the universe. Often we have a set of more than two events that partition the universe. For example, suppose we have n events $B_1, \cdots, B_n$ such that:

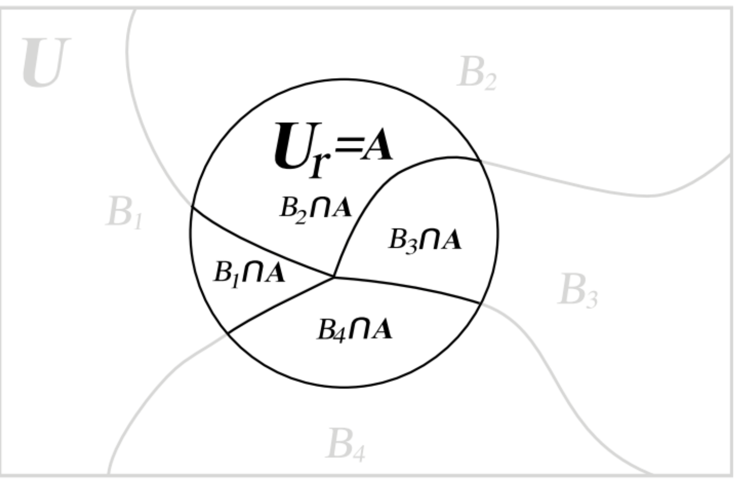
- The union $B_1 \cup B_2 \cup \cdots \cup B_n = U$, the universe, and
- Every distinct pair of the events are disjoint, $B_i \cap B_j = \phi$ for $i = 1, \ldots, n$, $j = 1, \ldots, n$, and $i \neq j$.

Then we say the set of events $B_1, \cdots, B_n$ *partitions* the universe. An observable event A will be partitioned into parts by the partition. $A = (A \cap B_1) \cup (A \cap B_2) \cup \ldots (A \cap B_n)$. $(A \cap B_i)$ and $(A \cap B_j)$ are disjoint since B_i and B_j are disjoint. Hence

$$\boxed{P(A) = \sum_{j=1}^{n} P(A \cap B_j)\,.}$$

Teorema Bayes (*Cont.*)





$$P(A) = \sum_{j=1}^{n} P(A \cap B_j)$$

$$P(A) = \sum_{j=1}^{n} P(A|B_j) \times P(B_j)$$

Teorema Bayes (*Cont.*)

$$P(A) = \sum_{j=1}^{n} P(A \cap B_j)$$

$$P(A) = \sum_{j=1}^{n} P(A|B_j) \times P(B_j)$$

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

Bayes' Rule :

$$\boxed{P(B_i|A) = \frac{P(A|B_i) \times P(B_i)}{\sum_{j=1}^{n} P(A|B_j) \times P(B_j)}}$$

Studi Kasus 1:

- The chance of a certain medical test being positive is 90%, if a patient has disease D. Known 1% of the population have the disease, and the test records a false positive 5% of the time. If you receive a positive test, what is your probability of having D?

- We are told: $P(+|D) = 0.90, \quad P(D) = 0.01, \quad P(+|\tilde{D}) = 0.05$

- We want to know: $P(D|+)$

- Bayes' Theorem: $P(D|+) = \frac{P(+|D)\,P(D)}{P(+)} = \frac{P(+|D)\,P(D)}{[P(+|D)\,P(D)]+[P(+|\tilde{D})\,P(\tilde{D})]}$

- Substituting in the data: $P(D|+) = \frac{0.9\times0.01}{(0.9\times0.01)+(0.05\times0.99)} = 0.15$

- *Interpretation: although the test is correct 90% of the time, the probability of having D after a positive test is only 15%. This is because only a small fraction of the population have the disease.*

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

Bayes' Theorem: The Key to Bayesian Statistics

To see how we can use Bayes' theorem to revise our beliefs on the basis of evidence, we need to look at each part. Let $B_1, \ldots, B_n$ be a set of unobservable events which partition the universe. We start with $P(B_i)$ for $i = 1, \ldots, n$, the *prior* probability for the events B_i, for $i = 1, \ldots, n$. This distribution gives the weight we attach to each of the B_i from our prior belief. Then we find that A has occurred.

The *likelihood* of the unobservable events $B_1, \ldots, B_n$ is the conditional probability that A has occurred given B_i for $i = 1, \ldots, n$. Thus the *likelihood* of event B_i is given by $P(A|B_i)$. We see the *likelihood* is a function defined on the events $B_1, \ldots, B_n$. The *likelihood* is the weight given to each of the B_i events given by the occurrence of A.

$P(B_i|A)$ for $i = 1, \ldots, n$ is the *posterior* probability of event B_i, given that event A has occurred. This distribution contains the weight we attach to each of the events B_i for $i = 1, \ldots n$ after we know event A has occurred. It combines our prior beliefs with the evidence given by the occurrence of event A.

Prior, Posterior, Likelihood, dan Marginal Likelihood

$$P(H|D) = \frac{P(H)P(D|H)}{P(D)}$$

In Bayesian statistics, most of the terms in Bayes' rule have special names. Some of them even have more than one name, with different scientific communities preferring different terminology. Here is a list of the various terms and the names we will use for them:

- $P(H|D)$ is the **posterior probability**. It describes how certain or confident we are that hypothesis H is true, given that we have observed data D. Calculating posterior probabilities is the main goal of Bayesian statistics!
- $P(H)$ is the **prior probability**, which describes how sure we were that H was true, before we observed the data D.
- $P(D|H)$ is the **likelihood**. If you were to assume that H is true, this is the probability that you would have observed data D.
- $P(D)$ is the **marginal likelihood**. This is the probability that you would have observed data D, *whether H is true or not.*

We often write Bayes' theorem in its proportional form as

$$posterior \propto prior \times likelihood$$

Prior Diskret

Studi Kasus 2:

(Teorema Bayes untuk Binomial dengan Prior Diskret)

Let $Y|\pi$ be $binomial(n = 4, \pi)$. Suppose we consider that there are only three possible values for π, .4,.5, and .6. We will assume they are equally likely. The joint probability distribution $f(\pi_i, y_j)$ is found by multiplying the conditional observation distribution $f(y_j|\pi_i)$ times the prior distribution $g(\pi_i)$. In this case, the conditional observation probabilities come from the $binomial(n = 4, \pi)$ distribution.

Suppose $Y = 3$ was observed.

Bayes' Box

π	*prior*	*likelihood*	*prior* $\times$ *likelihood*	*posterior*
.4	$\frac{1}{3}$	.1536	.0512	$\frac{.0512}{.2497} = .205$
.5	$\frac{1}{3}$	.2500	.0833	$\frac{.0833}{.2497} = .334$
.6	$\frac{1}{3}$	.3456	.1152	$\frac{.1152}{.2497} = .461$
marginal $P(Y = 3)$			.2497	1.000

Bayes' Box

π	*prior*	*likelihood*	*prior* $\times$ *likelihood*	*posterior*
.4	$\frac{1}{3}$	.1536	.0512	$\frac{.0512}{.2497} = .205$
.5	$\frac{1}{3}$	.2500	.0833	$\frac{.0833}{.2497} = .334$
.6	$\frac{1}{3}$	.3456	.1152	$\frac{.1152}{.2497} = .461$
marginal $P(Y = 3)$			.2497	1.000

- Put in the *parameter values*, the *prior*, and the *likelihood* in their respective columns. The *likelihood* values are *binomial(n,*π_i*)* evaluated at the observed value of y.

- Multiply each element in the *prior* column by the corresponding element in the *likelihood* column and put in the *prior* $\times$ *likelihood* column.

- Sum these *prior* $\times$ *likelihood*.

- Divide each element of *prior* $\times$ *likelihood* column by the sum of *prior* $\times$ *likelihood* column. (This rescales them to sum to 1.)

- Put these in the *posterior* column!

Studi Kasus 3:

(Teorema Bayes untuk Poisson dengan Prior Diskret)

Let $Y|\mu$ be Poisson(μ). Suppose that we believe there are only four possible values for μ, 1,1.5,2, and 2.5. Suppose we consider that the two middle values, 1.5 and 2, are twice as likely as the two end values 1 and 2.5. Suppose $y = 2$ was observed. Plug the value $y = 2$ into formula

$$f(y|\mu) = \frac{\mu^y e^{-\mu}}{y!}$$

to give the likelihood.

Bayes' Box

μ	*prior*	*likelihood*	*prior* $\times$ *likelihood*	*posterior*
1.0	$\frac{1}{6}$	$\frac{1.0^2 e^{-1.0}}{2!} = .1839$	.0307	$\frac{.0307}{.2473} = .124$
1.5	$\frac{1}{3}$	$\frac{1.5^2 e^{-1.5}}{2!} = .2510$	.0837	$\frac{.0837}{.2473} = .338$
2.0	$\frac{1}{3}$	$\frac{2.0^2 e^{-2.0}}{2!} = .2707$	.0902	$\frac{.0902}{.2473} = .365$
2.5	$\frac{1}{6}$	$\frac{2.5^2 e^{-2.5}}{2!} = .2565$	.0428	$\frac{.0428}{.2473} = .173$
marginal $P(Y=2)$			.2473	1.000

Materi Praktikum

1

Suppose there is a medical diagnostic test for a disease. The *sensitivity* of the test is .95. This means that if a person has the disease, the probability that the test gives a positive response is .95. The *specificity* of the test is .90. This means that if a person does not have the disease, the probability that the test gives a negative response is .90, or that the *false positive* rate of the test is .10. In the population, 1% of the people have the disease. What is the probability that a person tested has the disease, given the results of the test is positive? Let D be the event "the person has the disease" and let T be the event "the test gives a positive result."

2

Suppose there is a medical screening procedure for a specific cancer that has *sensitivity* = .90, and *specificity* = .95. Suppose the underlying rate of the cancer in the population is .001. Let B be the event "the person has that specific cancer," and let A be the event "the screening procedure gives a positive result."

(a) What is the probability that a person has the disease given the results of the screening is positive?

(b) Does this show that screening is effective in detecting this cancer?

3

There is an urn containing 9 balls, which can be either green or red. The number of red balls in the urn is not known.

Let X be the number of red balls in the urn.

Suppose we look at the two draws from the urn (without replacement) as a single experiment. The results were first draw red, second draw green. Find the posterior distribution of X by filling in the simplified table.

X	*prior*	*likelihood*	*prior* $\times$ *likelihood*	*posterior*

4

Selesaikan problem ini melalui Bayes' Box untuk memperoleh posterior: (a).Secara manual; (b).Menggunakan Program R.

Let Y_1 be the number of successes in $n = 10$ independent trials where each trial results in a success or failure, and π, the probability of success, remains constant over all trials. Suppose the 4 possible values of π are .20, .40, .60, and .80. We do not wish to favor any value over the others so we make them equally likely. We observe $Y_1 = 7$. Find the posterior distribution by filling in the simplified table.

π	*prior*	*likelihood*	*prior* $\times$ *likelihood*	*posterior*

Pustaka

1. Reich BJ dan Ghosh SK. (2019). *Bayesian Statistical Methods*. Taylor and Francis Group.

2. Bolstad WM dan Curran JM. (2017). *Introduction to Bayesian Statistics 3^th^ Edition*. John Wiley and Sons, Inc.

3. Lee, Peter M. (2012). *Bayesian Statistics, An Introduction 4^th^ Edition*. John Wiley and Sons, Inc.

4. Albert, Jim. (2009). *Bayesian Computation with R*. Springer Science and Business Media.

5. Ghosh JK, Delampady M, dan Samanta T. (2006). *An Introduction to Bayesian Analysis, Theory and Methods*. Springer Science and Business Media.

Terima Kasih