Part TWO

Mathematical Methods for Physics

from **On the Studies of Physics and Her Axillary Studies** by Shing Hin (John) Yeung

Chapter 35

Bayesian Inference

What is in this chapter?
Infering probabilities due to independent conditions can be done by conditional probability, and this is further elaborated through Bayesian inference. Bayesian inference is centralised due to Bayes Theorem and it is still emerging to more applications. In this chapter, basic terms will be explained and under the following order:

35.1 Conditional Probability Revisited

A conditional probability is a value of probability that is subjected to both:

- **Prior** where noted as event A in eq. (35.1)
- **Evidence** where noted as event B in eq. (35.1)

where the notation of conditional probability is

$$P(A \mid B) \tag{35.1}$$

where it reads as

> Probability of A due to B.

This is the basis of Bayesian statistics. Conditional probability is very useful in daily statements, such as

> How often I will detect α-particles from Americium-242, given that the background count is $n_{\text{back}} = 7$?

We can rewrite this statement as

$$P\left(\alpha\text{-particles from } {}^{242}_{95}\text{Am} \mid n_{\text{back}} = 7\right) \tag{35.2}$$

like in eq. (35.1). In this chapter we would be using conditional probability to justify the physical world, such as how some prior assumptions may affect our observed results.

In particular, we rather to use the equivalence of conditional probability eq. (35.1) as

$$P(A \mid B) = \frac{P(A \wedge B)}{P(B)} \tag{35.3}$$

in which $P(A \wedge B)$ means when events A and B is true or measured simultaneously. $P(B)$ which is the evidence of the posterior $P(A \mid B)$, is the marginal probability which is

$$P(B) = \int P(B, A)\, \mathrm{d}B \tag{35.4}$$

$$P(B) = \int P(A \mid B) \cdot P(A)\, \mathrm{d}B \tag{35.5}$$

where eq. (35.5) derives from eq. (35.2). Equations (35.4) and (35.5) means of normalising the posterior probability $P(A \mid B)$.

35.1.1 Expected Conditional Probability

An important relation is the expectation of the expectation value of the conditional variable is the expectation of the random variable. Which reads

$$\langle X \rangle = \langle \langle X \mid \theta \rangle \rangle \tag{35.6}$$

35.2 Bayes' Theorem

The central of Bayesian inference is due to this equation:

$$P(A \mid B) = \frac{P(B \mid A) \times P(A)}{P(B)} \tag{35.7}$$

and this notation brought down from eq. (35.1) in which we ought to know the probability of event A, due to the evidence event B is true. Equation (35.7) can be explained as

$$\text{Posterior} = \frac{\text{Likelihood} \times \text{Prior}}{\text{Evidence}} \tag{35.8}$$

where the concepts of prior is explained in section 35.2.1, likelihood is explained in section 40.3 in page 228.

35.2.1 Prior is the Summarised Probability

Prior distribution is key to Bayesian inference in which, it represents that data that has been observed. The posterior probability in eq. (35.7) updates the probability with new information provided. Gelman (2002) has setted up two points to consider when using prior distributions (Gelman 2002, p.1 from the excerpt):

- "What information is going into the prior distribution" So that the well-defined parameters and large sample size will minimise the effect of prior distributions onto the posterior distribution.

- "The properties of the resulting posterior distribution that affected by the prior" In particular the large sample size affect the posterior distribution less. Also, the data is preferred to reflect direct information of the parameters. For example, the dense data may explain where the mean value is located, and this can be achieved by large numbers of samples as well.

Hence setting up well-behaved prior distribution is very important in the analysis, but how? One approach may be done is to induct the probability, such as from the data that we already knew. For example, in a situation of a ball fell into one of the holes. You are standing under the whole, so as you know the ball has hit the ground, but not when it passes through any hole above you. Given that there are three holes above you. You concluded that the prior that means

> the possibility of how possible for a ball to pass through a hole is one-third (i.e. $\frac{1}{3}$).

35.3 Bayesian is a Subjective Probability

The philosophy of eqs. (35.7) and (35.8) means of prior probability maps onto the posterior probability. The prior which means a function of (solely by) the random variable of interest, means the subjects' belief.

35.4 There is no meaning if it is not Probability in the Bayesian Inference

35.5 Thomas Bayes never claimed Bayesian Statistics

Bayes' Theorem which discovered by Thomas Bayes, and his manuscript was significantly edited by Richard Price and read at the Royal Society. Sadly, Bayes was already dead two (i.e. 2) years ago and he could not see the very important applications due to his theorem.

Bibliography

Box, G. E. P. and G. C. Tiao (1973). **Bayesian inference in statistical analysis**. en. Addison-Wesley Pub. Co.

Gelman, A. (2002). **Prior distribution**. [Online; accessed 18-January-2017]. URL: http://www.stat.columbia.edu/~gelman/research/published/p039-_o.pdf.

Appendix 35.A Working Diary

35.A.1 16/01/2017

The chapter is one of the chapter originally here at start, I expanded this chapter today.

35.A.2 17/01/2017

Took a day off from uni. Make the acknowledgement for the code tmr.

35.A.3 18/02/2017

Nearly finished this chapter, thinking of exercises.

35.A.4 23/01/2017

Box and Tiao (1973) inspired me to finish section 35.3.

Appendix 35.B Selected Readings and their Reviews

Appendix 35.C Connections to Other Topics

Appendix 35.D Documentations