An Introduction to Bayesian Statistic: Embracing the Power of Probability

By Dr Nur Aazifah[1] Dr Hani Syahida[2] Associated Prof Dr Chew Boon How[1,3]

1) CRU HSAAS, UPM

2) Fakulti Sains, Jabatan Matematik dan Statistik, UPM

3) Fakulti Perubatan dan Sains Kesihatan, UPM

What is Bayesian Statistic:

It is a branch of statistical inferences that revolves around probability. The core of Bayesian statistics lies in Bayes' theorem which helps us to understand and utilise conditional probability. The conditional probability will become the basis of prior information. One famous use in the health science of the theorem is in the calculation of positive and negative predictive values. Readers are encouraged to read: What clinician should Know about Sensitivity, specificity and predictive value? It provides sufficient knowledge to understand the concept of the theorem.

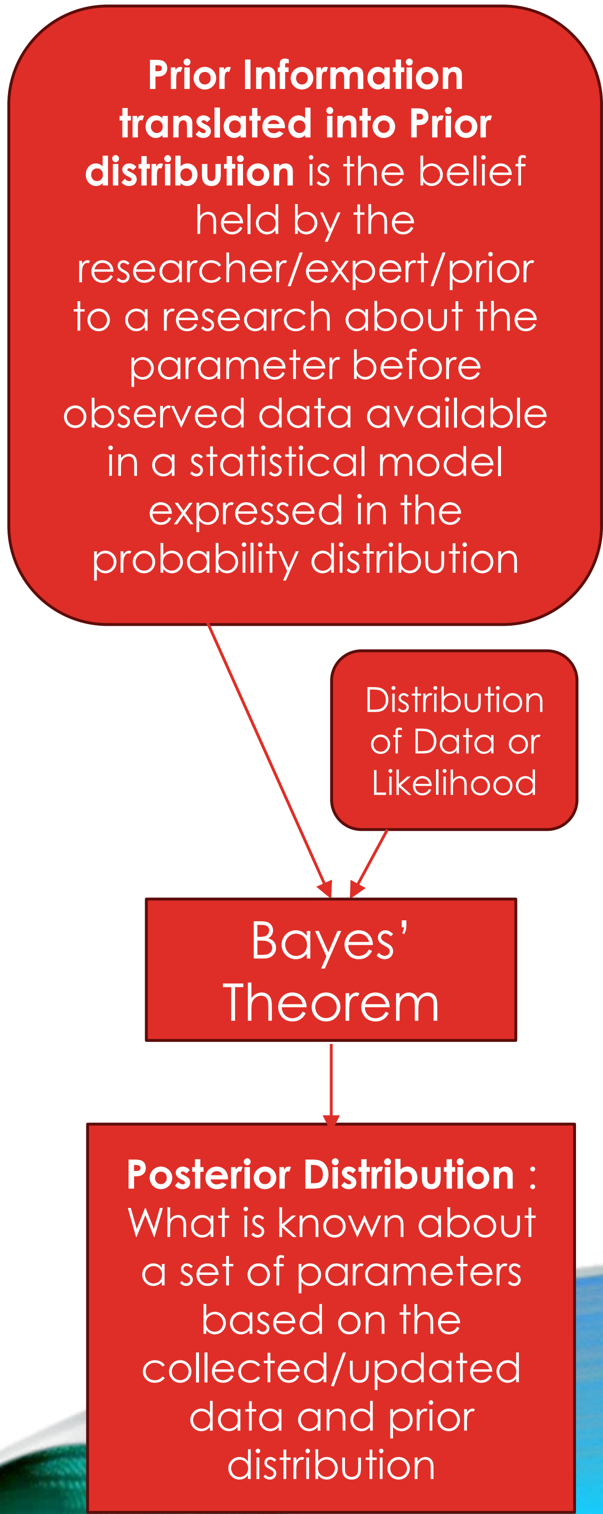
What is the goal of Bayesian Statistics?

To represent prior uncertainty about model parameters with a probability distribution and to update this prior uncertainty with observed/current data to produce a **posterior distribution** of the parameter with the belief it will produce less uncertainty. Inference from Bayesian analysis produces from the posterior distribution. Bayes' theorem for a probability distribution is often stated as :

Posterior distribution is proportional to Likelihood X Prior

What is the main difference between Bayesian statistics and Frequentist statistics?

In Bayesian, the parameter is treated as a random variable which will be translated as probability distribution. While frequentist statistic treats parameter as fixed quantities which means it has a constant value and is not subject to any uncertainty across different sample. Others different illustrated in the next section.



HOW BAYESIAN STATISTICS DIFFER FROM THE FREQUENTIST STATISTICS?

Aspect	Bayesian Statistics	Frequentist Statistic
Philosophical Approach	Subjective and incorporates prior belief.	Objective and does not use prior belief
Treatment to parameter and data	Treat parameter and data as random	Treat parameter as fixed while data is random.
Hypothesis testing	Bayes factor which is the ratio of the marginal likelihood of the observed data under the two hypothesis.	Relies on significant level or P-value
Sample size	Can be beneficial for small sample size	Use the Laws of Large numbers and Central limit theorem, ideally, large sample size is preferred. The minimal sample size requires typically depends on the minimal different effect sizes, level of significance , power and type of statistical analysis use.
Interpretation	Produces probability distributions for parameters,	Produced point estimates and confidence interval
Inferences	Make from posterior distribution	Make from likelihood
Prior information	Utilizes prior knowledge /belief about parameter usually define by hyperparameter.	Does not incorporate prior information. It solely relies on the observed data
Parameter estimation	Involve computing the posterior distribution of the parameter given the data and prior distribution.	Estimation commonly using likelihood estimation for non-normal distribution while for normal distribution usually using ordinary least square
Interval estimation	Credible intervals - Range of value of parameter that most likely to lie within interval that reflect the uncertainty in its estimation.	Confidence interval - Range of values base on repeated sampling from the same population that reflect the precision.
Likelihood usage	Incorporated data likelihood distribution to update belief into posterior distribution	Likelihood distribution used for estimation of parameter to form inferences
Computation	May involve complex computation especially with high dimensional parameter and complex distribution typically involve computational method Markov Chain Monte Carlo	Usually straight forward and computationally efficient
Decision making	Bayesian decision theory which involve Prior probability, likelihood function, Posterior probability, Bayes factor, Utility Function and Loss Function.	Usually using combination of point estimate, parameter, confident interval, clinical and statistically significant for decision making
Software	WinBUGS, JASP, BIEMS, AND SPSS, R and STATA in the latest version.	SPSS,SAS, R,STATA, Phyton

OPTIMAL SCENARIOS FOR BAYESIAN ANALYSIS: WHEN SHOULD WE USE IT?

In a complex models for examples :

- High-dimensional integration needed
- Multilevel latent variable model including those with random effect factor loadings, and random slopes when observe variables are categorical
- Three-level latent variable models that have categorical variable

Background knowledge can be incorporated in the analysis

- The notion of using prior research of other information and to produce updated prior is very reasonable. It will show the progression of the research toward a more refined knowledge.

Good for small sample size

- Bayesian analysis not based on large sample hence large sample is not a perquisite criteria for the math to work.
- With the advancement of software that able to integrated simulation the issue with sampling even for complex distribution can be solved.
- Many papers have shown the benefits of Bayesian statistic in the context of small data set. (Zhang et.al 2007)

REFERENCES:

1)Van de Schoot, R., Depaoli, S., King, R. *et al.* Bayesian statistics and modelling. *Nat Rev Methods Primers* 1, 1 (2021). https://doi.org/10.1038/s43586-020-00001-2

2)John W Stevems What Is Bayesian Statistic (2009)

3) Van de Schoot, R., & Depaoli, S. (2014). Bayesian analyses: where to start and what to report. *The European Health Psychologist, 16*(2), 75-84. Retrieved from http://www.ehps.net/ehp/index.php/contents/issue/view/ehp.v16.i2/showToc

4) Zhang et al (2007) Bayesian Analysis of Longitudinal data using growth curve model

EXAMPLE:

Click here

STATISTICAL MODELLING OF THE CONSISTENCY OF SYMPTOMS REPORTED DURING HYPOGLYCEMIA FOR INDIVIDUAL PATIENTS

- The research focuses on the development of Bayesian **latent variable statistical models** for evaluating the consistency of hypoglycemia symptoms in individual diabetes patients.
- The models in the paper are built using Bayesian methodology and Markov chain Monte Carlo techniques.
- Bayesian statistics is especially beneficial when dealing with complex distributions and limited sample sizes since it allows previous information and uncertainty to be incorporated into the analysis.
- Frequentist statistics, on the other hand, rely on high sample sizes and presume that the data is derived from a fixed distribution.
- As a result, the use of the Bayesian approach in this study is reasonable given the data's complicated distribution and small sample size.
- The findings of the study show the efficacy of Bayesian methodology in building statistical models for assessing the consistency of hypoglycemia patients.

References: HS Zulkifli Statistical Modelling of The consistency of symptoms reported during hypoglycemia for individual patient (2017)

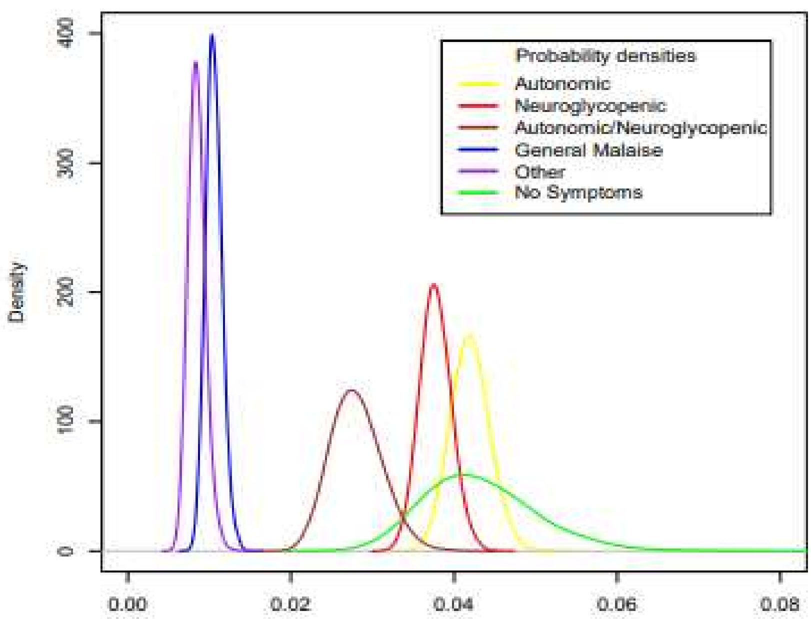


Figure 5.12: Posterior distributions of mean group propensity, u_l in hierarchical model.