ENGG2780 Statistics for Engineers

Ryan Chan

February 10, 2025

Abstract

This is a note for **ENGG2780 - Statistics for Engineers** for self-revision purpose ONLY. Some contents are taken from lecture notes and reference book.

Mistakes might be found. So please feel free to point out any mistakes.

Contents are adapted from the lecture notes of ENGG2760, prepared by Sinno Jialin Pan, as well as some online resources.

This course heavily relies on prior knowledge of probability (which you can refer to in the notes I wrote for ENGG2760). Therefore, before proceeding with this course, make sure you understand the foundation, as I will assume familiarity with probability concepts.

Contents

Chapter 1

Bayesian Statistic

1.1 Statistic v.s. Probability

Statistics focuses on real-life applications where the underlying distribution is often unknown. To address this, we use **statistical inference** to analyze observed data and estimate the unknown distribution. Rather than finding the exact distribution, we approximate it using models such as parametric (e.g., normal, exponential) or non-parametric approaches. Once a suitable model is chosen, probability laws help us make predictions and draw conclusions, though these approximations involve assumptions and uncertainties.

Now, let's move on to our first topic in statistics:

1.2 Bayesian Statistics

1.2.1 Introduction

In the probability course, we learned Bayes' Rule ENGG2760: Theorem 3.2.1, which helps us calculate conditional probabilities and, at times, update our beliefs based on new evidence.

And it turns out that one of the statistical inferences we use is based on Bayes' rule, namely Bayesian statistical inference. In Bayesian statistical inference, we: (1) assign prior probabilities to parameters; (2) observe data; and (3) update probabilities via Bayes' rule:

$$\underbrace{f_{\Theta|X}(\theta|x)}_{\text{Posterior}} = \frac{\overbrace{f_{\Theta}(\theta)}^{\text{Prior}}\overbrace{f_{X|\Theta}(x|\theta)}^{\text{Observation}}}{f_X(x)}$$

Here we have both the posterior and prior probabilities of the parameters θ and observations x.

We have four variations of the Bayes' rule shown above.

Condition	Bayes' rule
Θ discrete, X discrete	$p_{\Theta|X}(\theta|x) = \dfrac{p_{\Theta}(\theta)p_{X|\Theta}(x|\theta)}{\sum_{\theta'} p_{\Theta}(\theta')p_{X|\Theta}(x|\theta')}$
Θ discrete, X continuous	$p_{\Theta|X}(\theta|x) = \dfrac{p_{\Theta}(\theta)f_{X|\Theta}(x|\theta)}{\sum_{\theta'} p_{\Theta}(\theta')f_{X|\Theta}(x|\theta')}$
Θ continuous, X discrete	$f_{\Theta|X}(\theta|x) = \dfrac{f_{\Theta}(\theta)p_{X|\Theta}(x|\theta)}{\int f_{\Theta}(\theta')p_{X|\Theta}(x|\theta')}$
Θ continuous, X continuous	$f_{\Theta|X}(\theta|x) = \dfrac{f_{\Theta}(\theta)f_{X|\Theta}(x|\theta)}{\int f_{\Theta}(\theta')f_{X|\Theta}(x|\theta')}$

We can use $Z(x)$ to denote the denominator for both discrete and continuous cases. It depends only on the observed data x.

> **Example** (Probability Review). We flip a coin. How likely is it to get 2 heads in 3 coin flips if the probability of heads is p, where p could be 0.5, 0.7, and 1?
>
> Also, use the Central Limit Theorem to estimate the probability of at least 200 heads in 300 coin flips.
>
> **Solution:**
> $$\mathbb{P}(H = 2) = \binom{3}{2} p^2 (1 - p)$$
>
> $p = 0.5 : \mathbb{P}(H = 2) = \binom{3}{2} \times 0.5^2 \times 0.5 = 0.375$
>
> $p = 0.7 : \mathbb{P}(H = 2) = \binom{3}{2} \times 0.7^2 \times 0.3 = 0.441$
>
> $p = 1 : \mathbb{P}(H = 2) = \binom{3}{2} \times 1^2 \times 0 = 0$
>
> For the probability of at least 200 heads in 300 coin-flips,
> $$H \sim \text{Binomial}(300, p), \quad \mu = 300p, \quad \sigma = \sqrt{300p(1 - p)}$$
>
> $p = 0.5 : \mu = 150, \sigma = 8.66$
> $$\begin{aligned} \mathbb{P}(H \geq 200) &= \mathbb{P}(\frac{H - 150}{8.66} \geq \frac{200 - 150}{8.66}) \\ &= \mathbb{P}(z \geq 5.77) \\ &\approx 0 \end{aligned}$$
>
> $p = 0.7 : \mu = 210, \sigma = 7.94$
> $$\begin{aligned} \mathbb{P}(H \geq 200) &= \mathbb{P}(\frac{H - 210}{7.94} \geq \frac{200 - 210}{7.94}) \\ &= \mathbb{P}(z \geq -1.26) \\ &= \Phi(1.26) \\ &= 0.896 \end{aligned}$$

Again, we flip a coin three times and get two heads. You are told that there are three types of coins with different priors, but you don't know which coin you are flipping. It is obvious that the first coin flip will affect your belief (prior) about which coin you have. For example, if you see 100 heads out of 100 flips, you might strongly believe that both sides of the coin are heads. But to what extent does each flip influence your belief? This brings us to the problem of statistics.

> **Example.** A coin can be one of three types:
>
> 1. A fair coin $\theta = 1$ with one head and one tail – 90%
>
> 2. A coin $\theta = 2$ with both sides as heads – 5%
>
> 3. A coin $\theta = 3$ with both sides as tails – 5%
>
> Now, you flip a head without knowing which coin you have. How should you update your belief (priors)?
>
> **Solution:**
> $$\mathbb{P}(\theta = 1 | H_1) = \frac{\mathbb{P}(H_1 | \theta = 1) \mathbb{P}(\theta = 1)}{Z(H_1)} = \frac{0.5 \times 0.9}{Z(H_1)} = \frac{0.45}{Z(H_1)}$$
>
> $$\mathbb{P}(\theta = 2 | H_1) = \frac{\mathbb{P}(H_1 | \theta = 2) \mathbb{P}(\theta = 2)}{Z(H_1)} = \frac{1 \times 0.05}{Z(H_1)} = \frac{0.05}{Z(H_1)}$$
>
> $$\mathbb{P}(\theta = 3 | H_1) = 0$$

Then we have $\mathbb{P}(H_1) = Z(H_1) = 0.45 + 0.05 + 0 = 0.5$

$$\mathbb{P}(\theta = 1|H_1) = \frac{0.45}{Z(H_1)} = 0.9 \quad \mathbb{P}(\theta = 1|H_1) = \frac{0.05}{Z(H_1)} = 0.1 \quad \mathbb{P}(\theta = 1|H_1) = 0$$

From this, we can update our belief, which we can then use to further readjust our belief if the second flip also results in a head.

$$\mathbb{P}(\theta = 1|H_2 H_1) = \frac{\mathbb{P}(H_2|\theta = 1, H_1)\mathbb{P}(\theta = 1|H_1)}{Z(H_2, H_1)} = \frac{0.5 \times 0.9}{Z(H_2, H_1)} = \frac{0.45}{Z(H_2, H_1)}$$

$$\mathbb{P}(\theta = 2|H_2 H_1) = \frac{\mathbb{P}(H_2|\theta = 2, H_1)\mathbb{P}(\theta = 2|H_1)}{Z(H_2, H_1)} = \frac{1 \times 0.1}{Z(H_2, H_1)} = \frac{0.1}{Z(H_2, H_1)}$$

$$\mathbb{P}(\theta = 3|H_2 H_1) = 0$$

Then we have $\mathbb{P}(H_2 H_1) = Z(H_2 H_1) = 0.45 + 0.01 + 0 = 0.55$

$$\mathbb{P}(\theta = 1|H_2 H_1) = \frac{0.45}{Z(H_2 H_1)} = 0.82 \quad \mathbb{P}(\theta = 1|H_2 H_1) = \frac{0.1}{Z(H_2 H_1)} = 0.18 \quad \mathbb{P}(\theta = 1|H_2 H_1) = 0$$

We also have Bayes's rule for multiple random variables:

$$f_{\Theta|X_1,\cdots,X_n}(\theta|x_1,\cdots,x_n) = \frac{f_{X_1,\cdots,X_n|\Theta}(x_1,\cdots,x_n|\theta)f_\Theta(\theta)}{Z(x_1,\cdots,x_n)}$$
$$\propto f_{X_1,\cdots,X_n|\Theta}(x_1,\cdots,x_n|\theta)f_\Theta(\theta)$$
$$= \underbrace{f_{X_1|\Theta}(x_1|\theta)\cdots f_{X_n|\Theta(x_n|\theta)}}_{\text{product of likelihood}}\underbrace{f_\Theta(\theta)}_{\text{prior}}$$

if $X_1, \cdots, X_n$ are independent given Θ.

1.2.2 Bayesian Statistical Inference

For Bayesian statistics, we have only one formula: Bayes's rule:

$$\underbrace{f_{\Theta|X}(\theta|x)}_{\text{posterior}} \propto \underbrace{f_{X|\Theta}(x|\theta)}_{\text{likelihood}}\underbrace{f_\Theta(\theta)}_{\text{prior}}$$

We have some prior knowledge, and after observing something, we can use the prior and likelihood to update our belief, which gives us the posterior. This posterior can later serve as the prior for another observation, allowing us to continuously update our belief throughout the observation process.

Example. Romeo is waiting for Juliet on their first date. He wants to estimate how long he will have to wait for her. Given that Romeo has some prior dating experience, he already has some prior knowledge about how late girls tend to be.

Girl A - $X \sim \text{Uniform}(0, 0.3)$;

Girl B - $X \sim \text{Uniform}(0, 0.8)$;

Girl C - $X \sim \text{Uniform}(0, 0.6)$,

where the uniform random variable shows the range of lateness. For example, for girl A, she will be late between the dating time and the dating time plus 0.3 hours. Then, how could you use Bayesian statistics to estimate the waiting time for Romeo's new girlfriend?

Solution: Here we can set up the uniform random variable $\text{Uniform}(0, \Theta)$, where Θ depends on the girls' decision. Then what we need to find is the θ for Juliet. We can then have

$$f_{X|\Theta}(x|\theta) = \begin{cases} \dfrac{1}{\theta}, & \text{if } 0 \le x \le \theta; \\ 0, & \text{otherwise.} \end{cases}$$

In Romeo's model, θ is also a uniform random variable $\theta \sim \text{Uniform}(0,1)$, where $X \sim \text{Uniform}(0, \Theta)$. Given that on their first date, Juliet arrived $\frac{1}{2}$ hours late, we have

$$f_{\Theta|X}(\theta|\tfrac{1}{2}) \propto f_\Theta(\theta) f_{X|\Theta}(\tfrac{1}{2}|\theta) = \frac{1}{\theta}$$

Here we have the prior $f_\Theta(\theta) = 1$ if $0 \le \theta \le 1$, and the likelihood $f_{X|\Theta}(\frac{1}{2}|\theta) = \frac{1}{\theta}$ if $\frac{1}{2} \le \theta \le 1$. Keep in mind that the prior comes from Romeo's model, where he has never dated a girl who is late for more than 1 hour, and it may not be valid if $\theta > 1$, which shows the limitation of Bayesian statistics. Also, the observation (likelihood) shows the probability of Juliet arriving precisely at (or within a very small interval around) time plus 0.5. Therefore, we have $\theta \ge \frac{1}{2}$.

For the integral to be equal to 1, we need to find the constant term. This can be found using calculus:

$$\int_{\frac{1}{2}}^1 \frac{1}{\theta} d\theta = \ln \theta \Big|_{\frac{1}{2}}^1 = \ln 2 \implies f_{\Theta|X}(\theta|\tfrac{1}{2}) = \frac{1}{\theta \ln 2}$$

Here we have $\theta < \frac{1}{2} = 0$ because from the data, we know that $\theta \ge \frac{1}{2}$, which means the lateness parameter is at least $\frac{1}{2}$, so it is not possible for Juliet to arrive between the dating time and dating time plus 0.5. We also have $\theta > 1 = 0$ because from Romeo's prior knowledge, he knows that a girl would not be later than 1 hour.
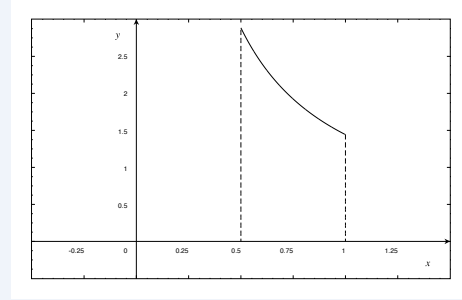
On their second date, Juliet arrived $\frac{1}{4}$ hours late. We then need to readjust the prior based on the previous model to find the new posterior.

$$f_{\Theta|X_1,X_2}\left(\theta\Big|\frac{1}{2},\frac{1}{4}\right) \propto f_{\Theta|X_1}\left(\theta\Big|\frac{1}{2}\right) f_{X_2|\Theta,X_1}\left(\frac{1}{4}\Big|\theta,\frac{1}{2}\right)$$

Here, since X_1 and X_2 are independent, we can discard X_1 in the calculation.

$$f_{\Theta|X_1,X_2}\left(\theta\Big|\frac{1}{2},\frac{1}{4}\right) \propto f_{\Theta|X_1}\left(\theta\Big|\frac{1}{2}\right) f_{X_2|\Theta}\left(\frac{1}{4}\Big|\theta\right) = \frac{1}{\theta \ln 2} \times \frac{1}{\theta} = \frac{1}{\theta^2 \ln(2)} \propto \frac{1}{\theta^2}$$

The same as above, we have $f_{X_2|\Theta}(\frac{1}{4}|\theta) = \frac{1}{\theta}$ since it is not possible for the lateness to be less than $\frac{1}{4}$ hours. Also, given the prior as calculated in the first part, we have $f_{\Theta|X_1}(\theta|\frac{1}{2}) = \frac{1}{\theta \ln 2}$ if $\frac{1}{2} \le \theta \le 1$.

For the integral to be equal to 1, we need to find the constant term. This can be found using calculus:

$$\int_{\frac{1}{2}}^1 \frac{1}{\theta^2} d\theta = 1 \implies f_{\Theta|X_1,X_2}\left(\theta\Big|\frac{1}{2},\frac{1}{4}\right) = \frac{1}{\theta^2}$$
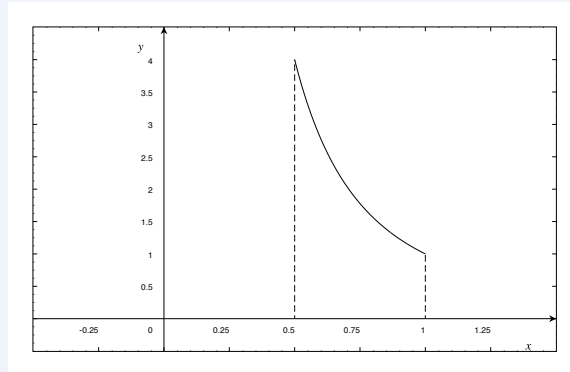
Remark (Bayes' rule variant).

$$\mathbb{P}(\theta|x_1, x_2) = \frac{\mathbb{P}(x_2|\theta, x_1)\mathbb{P}(\theta|x_1)}{\mathbb{P}(x_2|x_1)}$$

Proof.

$$f_{\Theta|X_1,X_2}\left(\theta\Big|\frac{1}{2}, \frac{1}{4}\right) = \frac{f_{\Theta,X_1,X_2}\left(\theta, \frac{1}{2}, \frac{1}{4}\right)}{f_{X_1,X_2}\left(\frac{1}{2}, \frac{1}{4}\right)}$$

$$= \frac{f_{X_2|\Theta,X_1}\left(\frac{1}{4}\Big|\theta, \frac{1}{2}\right) f_{\Theta,X_1}\left(\theta, \frac{1}{2}\right)}{f_{X_1,X_2}\left(\frac{1}{2}, \frac{1}{4}\right)}$$

$$= \frac{f_{X_2|\Theta,X_1}\left(\frac{1}{4}\Big|\theta, \frac{1}{2}\right) f_{\Theta|X_1}\left(\theta\Big|\frac{1}{2}\right) f_{X_1}\left(\frac{1}{2}\right)}{f_{X_1,X_2}\left(\frac{1}{2}, \frac{1}{4}\right)}$$

$$= \frac{f_{X_2|\Theta,X_1}\left(\frac{1}{4}\Big|\theta, \frac{1}{2}\right) f_{\Theta|X_1}\left(\theta\Big|\frac{1}{2}\right)}{f_{X_2|X_1}\left(\frac{1}{4}\Big|\frac{1}{2}\right)}$$

Thus,

$$f_{\Theta|X_1,X_2}\left(\theta\Big|\frac{1}{2}, \frac{1}{4}\right) \propto f_{X_2|\Theta,X_1}\left(\frac{1}{4}\Big|\theta, \frac{1}{2}\right) f_{\Theta|X_1}\left(\theta\Big|\frac{1}{2}\right)$$

∎

Chapter 2

Known Bugs

Appendix A

Z TABLE

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990