Bayesian Computing Question Bank

Chapter 1

1. Define Bayesian Thinking. Compare Bayesian Thinking with Critical Thinking.
   Solution:

- Bayesian Thinking is a mental model that allows you to adapt your thinking reactively to new evidence.
- Bayesian thinking is a form of statistical reasoning. It involves calculating and updating probabilities as new information becomes available to make the best possible predictions.

Aspect	**Frequentist Approach**	**Bayesian Approach**
Probability Interpretation	Objective: Probabilities represent long-term frequencies or limiting behavior of repeated experiments.	Subjective: Probabilities represent degrees of belief or uncertainty based on prior knowledge and data.
Treatment of Parameters	Fixed: Parameters are fixed, unknown constants. Estimation involves finding the “best” estimate based on data.	Random: Parameters are treated as random variables with their own probability distributions. They are updated based on prior beliefs and data, resulting in posterior distributions.
Prior Information	N/A: Typically, prior information is not explicitly incorporated into the analysis.	Crucial: Bayesian analysis involves specifying prior distributions representing prior beliefs about parameters before observing data.
Inference Approach	Hypothesis Testing: Involves p-values and rejection regions.	Credible Intervals: Involves credible intervals to estimate parameter values with specified probabilities.
Handling Uncertainty	Point Estimates: Point estimates (e.g., sample mean) with associated uncertainties (e.g., confidence intervals).	Probability Distributions: Posterior distributions that directly model the uncertainty of parameter estimates.
Sample Size Requirement	Large Sample: Often requires a large sample size for accurate parameter estimation.	Smaller Sample: Bayesian methods can provide reasonable estimates even with smaller sample sizes, especially with informative priors.
Computational Complexity	Simpler: Often involves direct formulas for parameter	More Complex: Requires numerical methods like MCMC for

Aspect	Frequentist Approach	Bayesian Approach
	estimation (e.g., maximum likelihood).	posterior estimation, especially for complex models.
Hypothesis Testing	p-values and hypothesis tests are prone to misinterpretation and controversies.	Bayesian hypothesis testing uses Bayes Factors or posterior probabilities for direct comparison.
Model Selection	Relies on criteria like AIC or BIC.	Model comparison using posterior model probabilities (Bayes Factors) or marginal likelihoods.
Interpretation of Results	Focused on the data and observed effects.	Results interpreted in the context of prior beliefs and their update based on data.

2. Discuss Bayesian Robustness along with its key points.
   Solution:

- Bayesian robustness refers to the property of Bayesian statistical methods to provide stable and reliable inference even when the assumptions of the model are violated to some extent.
- In other words, Bayesian methods are designed to handle uncertainty in model assumptions and data by incorporating prior information, which can lead to more stable and credible results compared to some frequentist methods that might be sensitive to violations of assumptions.

Key points about Bayesian robustness

- **Prior Distribution:** Bayesian analysis involves specifying a prior distribution that reflects your beliefs or information about the parameters before observing the data. A strong prior can help regularize the estimates and make them more robust to outliers and deviations from assumptions.

- **Posterior Distribution:** The Bayesian approach updates the prior beliefs with observed data to obtain the posterior distribution of the parameters. This posterior distribution takes into account both the data and the prior, allowing for a more nuanced understanding of uncertainty.

- **Small Sample Sizes:** Bayesian methods can perform well with small sample sizes because the prior information can play a substantial role in shaping the posterior distribution, which helps stabilize the estimates.

- **Model Misspecification:** Bayesian methods can handle certain types of model misspecification because they integrate over the uncertainty in the model space. If the true data-generating process is different from the assumed model, Bayesian methods can still provide reasonable estimates by incorporating prior information.

- **Robust Priors:** Robust Bayesian priors are chosen to be less sensitive to extreme values or outliers. These priors help mitigate the influence of unusual observations on the parameter estimates.
- **Sensitivity Analysis:** Bayesian robustness can be assessed through sensitivity analysis. This involves examining how the results change when different prior distributions are used or when assumptions are altered.
- **Hierarchical Models:** Hierarchical Bayesian models are particularly useful for robust analysis. They allow for information to be borrowed from related groups or populations, leading to more stable parameter estimates.
- **Markov Chain Monte Carlo (MCMC):** MCMC methods, commonly used in Bayesian analysis, enable the exploration of complex posterior distributions, which is useful when dealing with robustness issues.

3. The incidence of a disease in the population is 1%. A medical test for the disease is 90% accurate in the sense that it produces a false reading 10% of the time, both: (a) when the test is applied to a person with the disease; and (b) when the test is applied to a person without the disease. A person is randomly selected from the population and given the test. The test result is positive (i.e. it indicates that the person has the disease). What is the probability that the person actually has the disease?
Solution:

Let's use notation established above for hypotheses and data: let $\mathcal{H}_+$ be the hypothesis (event) that the person has the disease and let $\mathcal{H}_-$ be the hypothesis they do not. Likewise, let $\mathcal{T}_+$ and $\mathcal{T}_-$ represent the data of a positive and negative screening test respectively. We are asked to compute $P(\mathcal{H}_+|\mathcal{T}_+)$.

We are given

$$P(\mathcal{T}_+|\mathcal{H}_+) = 0.99, \quad P(\mathcal{T}_+|\mathcal{H}_-) = 0.02, \quad P(\mathcal{H}_+) = 0.005.$$

From these we can compute the false negative and true negative rates:

$$P(\mathcal{T}_-|\mathcal{H}_+) = 0.01, \quad P(\mathcal{T}_-|\mathcal{H}_-) = 0.98$$

Bayes' theorem yields

$$P(\mathcal{H}_+|\mathcal{T}_+) = \frac{P(\mathcal{T}_+|\mathcal{H}_+)P(\mathcal{H}_+)}{P(\mathcal{T}_+)} = \frac{0.99 \cdot 0.005}{0.99 \cdot 0.005 + 0.02 \cdot 0.995} = 0.19920 \approx 20\%$$

4. Consider six loaded dice with the following properties. Die A has probability 0.1 of coming up 6, each of Dice B and C has probability 0.2 of coming up 6, and each of Dice D, E and F has probability 0.3 of coming up 6. A die is chosen randomly from the six dice and rolled twice. On both occasions, 6 comes up. What is the posterior probability distribution of θ , the probability of 6 coming up on the chosen die.
Solution:

Let y be the number of times that 6 comes up on the two rolls of the chosen die, and let θ be the probability of 6 coming up on a single roll of that die.
Then the Bayesian model is:

$$(y \mid \theta) \sim Bin(2,\theta)$$

$$f(\theta) = \begin{cases} 1/6, & \theta = 0.1 \\ 2/6, & \theta = 0.2 \\ 3/6, & \theta = 0.3. \end{cases}$$

In this case $y = 2$ and so

$$f(y \mid \theta) = \binom{2}{y}\theta^{y}(1-\theta)^{2-y} = \binom{2}{2}\theta^{2}(1-\theta)^{2-2} = \theta^{2}.$$

So $f(y) = \sum_{\theta} f(\theta) f(y \mid \theta) = \frac{1}{6}(0.1)^2 + \frac{2}{6}(0.2)^2 + \frac{3}{6}(0.3)^2 = 0.06.$

So $f(\theta \mid y) = \dfrac{f(\theta)f(y \mid \theta)}{f(y)} = \begin{cases} (1/6)0.1^2/0.06 = 0.02778, & \theta = 0.1 \\ (2/6)0.2^2/0.06 = 0.22222, & \theta = 0.2 \\ (3/6)0.3^2/0.06 = 0.75, & \theta = 0.3. \end{cases}$

5. A credit card company wants to determine the mean income of its card holders. It also wants to find out if there are any differences in mean income between males and females. A random sample of 225 male card holders and 190 female card holders was drawn, and the following results were obtained:

	Mean	Standard Deviation
Males	\$ 16450	\$3675
Females	\$ 13220	\$ 3050

Calculate the 95% confidence intervals for the mean income for males and females. Is there any evidence to suggest that, on average, males' and females' income differ? If so, describe the difference.

Solution:

95% confidence interval for male income

The true population variance, σ^2, is unknown, we follow the t–distribution, i.e.

$$\bar{x} \pm t_{p/2} \times \sqrt{s^2/n}.$$

Here,

$$\begin{aligned} \bar{x} &= 16450, \\ s^2 &= 3675^2 = 13505625 \qquad \text{and} \\ n &= 225. \end{aligned}$$

The value $t_{p/2}$ must be found from table 1.1. Recall that the degrees of freedom, $\nu = n-1$, and so here we have $\nu = 225 - 1 = 224$. Notice that table 1.1 only gives value of ν up to 29; for higher values, we use the ∞ row. Since we require a 95% confidence interval, we read down the 5% column, giving a t value of 1.96 (recall that this is the same as the value used if σ^2 were *known* and we used the normal distribution – that's because the t–distribution converges to the normal distribution as the sample size increases). Thus, the 95% confidence interval for μ is found as

$$\begin{aligned} 16450 &\pm 1.96 \times \sqrt{13505625/225}, \quad \text{i.e.} \\ 16450 &\pm 480.2. \end{aligned}$$

So, the 95% confidence interval is (£15969.80, £16930.20).

95% confidence interval for female income

Again, the true population variance, σ^2, is unknown, so again we use the t–distribution as

$$\bar{x} \pm t_{p/2} \times \sqrt{s^2/n}.$$

Now,

$$\begin{aligned} \bar{x} &= 13220, \\ s^2 &= 3050^2 \\ &= 9302500, \quad \text{and} \\ n &= 190. \end{aligned}$$

Again, since the sample size is large, we use the ∞ row of table 1.1 to obtain the value of $t_{p/2}$, and so the 95% confidence interval for μ is found as

$$\begin{aligned} 13220 &\pm 1.96 \times \sqrt{9302500/190}, \quad \text{i.e.} \\ 13220 &\pm 1.96 \times 221.27, \quad \text{i.e.} \\ 13220 &\pm 433.69. \end{aligned}$$

So, the 95% confidence interval is (£12786.31, £13653.69).

Since the 95% confidence intervals for males and females *do not overlap*, there *is* evidence to suggest that males' and females' incomes, on average, are different. Further, it appears that male card holders earn more than women.

6. A sample of size 15 is taken from a larger population; the sample mean is calculated as 12 and the sample variance as 25. What is the 95% confidence interval for the population mean μ?
   Solution:

We know that the confidence interval is given by

$$\bar{x} \;\pm\; t_{p/2}\sqrt{s^2/n},$$

where

$$\begin{aligned} n &= 15, \\ \nu &= n-1 = 15-1 = 14, \\ p &= 5\%, \\ \bar{x} &= 12 \qquad \text{and} \\ s^2 &= 25. \end{aligned}$$

We can find our t value by looking in the $p = 5\%$ column and the $\nu = 14$ row, giving a value of 2.145. Putting what we know into our expression, we get

$$\begin{aligned} 12 &\pm t_{2.5\%}\sqrt{\frac{25}{15}} \\ 12 &\pm 2.145\sqrt{\frac{25}{15}} \qquad \text{i.e.} \\ 12 &\pm 2.77. \end{aligned}$$

Hence, the confidence interval is $(9.23, 14.77)$.

Chapter 2

1. What is over-dispersion? How can the Beta-Binomial distribution model be used for over-dispersion?

Solution:

Over-dispersion is a phenomenon where the variance of observed data is larger than what would be expected from a theoretical distribution, like the Binomial distribution. The Beta-Binomial distribution is a probability distribution used to model over-dispersion in data. Over-dispersion occurs when the observed variation in a dataset is greater than what would be expected based on a simple Binomial distribution. This phenomenon often arises in real-world data, where factors such as unaccounted for heterogeneity or extra sources of variation lead to increased variability.

Here's how the Beta-Binomial distribution can be used to address over-dispersion:

(a) **Understanding Over-Dispersion:**
- In a standard Binomial distribution, we have two parameters: n (the number of trials) and p (the probability of success in each trial).
- The variance of a Binomial distribution is np(1-p). When the observed variance in the data is significantly larger than np(1-p), over-dispersion is present.

(b) **Introduction of Extra Parameters:**
- The Beta-Binomial distribution introduces two additional parameters, alpha (α) and beta (β), to the model.
- Alpha (α) represents the number of successful outcomes (i.e., "successes") in the dataset.
- Beta (β) represents the number of unsuccessful outcomes (i.e., "failures") in the dataset.

(c) **Adjusting Variability:**
- The extra parameters, alpha and beta, allow for a more flexible modeling of the variance.
- By adjusting the values of alpha and beta, the Beta-Binomial distribution can capture the extra variability seen in over-dispersed data.
- When alpha and beta are equal, the distribution becomes equivalent to the standard Binomial distribution. However, by setting them differently, you can increase or decrease the variance as needed.

(d) **Fitting the Model:**
- To use the Beta-Binomial distribution for over-dispersion, you estimate the parameters alpha and beta from your data.
- This can often be done using statistical software or algorithms that perform maximum likelihood estimation or Bayesian inference.
- Once you've estimated alpha and beta, you can use the Beta-Binomial distribution to generate probability distributions that better match your observed data.

(e) **Applications:**
- The Beta-Binomial distribution is commonly employed in various fields, including biology, epidemiology, finance, and quality control, where data often exhibit over-dispersion due to multiple sources of variation.

2. For Beta distribution with shape parameter α, scale parameter β, mean μ and standard deviation σ, prove that $\alpha = \frac{\mu^2 - \mu^3}{\sigma^2} - \mu$.

Solution:

For Beta distribution, mean $\mu = \frac{\alpha}{\alpha+\beta}$ --------------------(1)

and $\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$ ------------------------------------(2)

From (1), $\alpha = \mu\alpha + \mu\beta \quad or\ \beta = \frac{1-\mu}{\mu}\alpha$

Substitute this in equation (2)

$$\sigma^2 = \frac{\frac{1-\mu}{\mu}\alpha^2}{\left(\alpha+\frac{1-\mu}{\mu}\alpha+1\right)\left(\alpha+\frac{1-\mu}{\mu}\alpha\right)^2}$$

$$= \frac{\frac{1-\mu}{\mu}\alpha^2}{\left(\frac{\mu\alpha+\alpha-\alpha\mu+\mu}{\mu}\right)\left(\frac{\mu\alpha+\alpha-\mu\alpha}{\mu}\right)^2}$$

$$= \frac{\left(\frac{1-\mu}{\mu}\right)\alpha^2}{\left(\frac{\alpha+\mu}{\mu}\right)\left(\frac{\alpha^2}{\mu^2}\right)}$$

$$\therefore \frac{1}{\mu^2}\sigma^2 = \frac{1-\mu}{\alpha+\mu}$$

$$\therefore \alpha + \mu = \frac{1-\mu}{\frac{1}{\mu^2}\sigma^2} = \frac{\mu^2-\mu^3}{\sigma^2}$$