

Bayesian Statistics: Sample Exam Questions

Abstract

This document contains sample exam questions and answers for a Bayesian Statistics course, covering topics such as maximum likelihood estimation, confidence intervals, likelihood functions, Bayesian updating, credible intervals, and posterior distributions.

December 1, 2025

Contents

1	Sample Bayesian Statistics Exam Questions	3
1.1	Question 1 — MLE for p in a coin flip	3
1.2	Question 2 — Lower end of a 95% CI for p	3
1.3	Question 3 — Upper end of a 95% CI for p	4
1.4	Question 4 — Definition of the likelihood	4
1.5	Question 5 — MLE for λ in an Exponential model	4
1.6	Question 6 — MLE for μ in a Normal model	5
2	Exam 2	6
2.1	Questions 1–5: Political Preferences	6
2.1.1	Question 1	6
2.1.2	Question 2	6
2.1.3	Question 3	7
2.1.4	Question 4	7
2.1.5	Question 5	8
2.2	Questions 6–9: Loaded Coins	8
2.2.1	Question 6	8
2.2.2	Question 7	9
2.2.3	Question 8	9
2.2.4	Question 9	10
3	Continuous Bayes, Credible Intervals	11
3.1	Question 1 — When to use the continuous Bayes formula	11
3.2	Question 2 — Sequential update after two heads	11
3.3	Question 3 — Interpretation of a prior probability	12
3.4	Question 4 — Interpretation of a posterior probability	13
3.5	Question 5 — Quantiles for a 90% equal-tailed interval	13
3.6	Question 6 — 95% equal-tailed interval for $N(0, 1)$	14
3.7	Question 7 — Meaning of HPD	14
3.8	Question 8 — 50% credible interval	14
4	Bayesian vs frequentist	16
4.1	Basic Interpretations	16
4.2	When to Prefer Frequentist Confidence Intervals	16
4.2.1	(a) Regulatory or standardized settings	16
4.2.2	(b) Need for guaranteed coverage properties	17
4.2.3	(c) Avoiding or lacking a prior	17
4.2.4	(d) Sample size and design calculations	17
4.3	When to Prefer Bayesian Credible Intervals	17
4.3.1	(a) Direct probability statements about parameters	17

4.3.2	(b) Presence of genuine prior information	17
4.3.3	(c) Complex or hierarchical models	18
4.3.4	(d) Decision-making with losses or utilities	18
4.4	When They Coincide (Approximately)	18
4.5	Compact Summary	18
5	Honor Quiz	20
5.1	Question 1	20
5.2	Question 2	20
5.3	Question 3	21
5.4	Question 4	22

1 Sample Bayesian Statistics Exam Questions

1.1 Question 1 — MLE for p in a coin flip

Question. In 100 flips of a coin, you observe 47 heads and 53 tails. Report the MLE \hat{p} of the probability of obtaining heads.

Answer.

$$\hat{p} = 0.47.$$

Explanation. For a Bernoulli (or Binomial) model, the likelihood for p based on n trials and x heads is

$$L(p \mid x) \propto p^x (1 - p)^{n-x}.$$

Maximizing this over p gives

$$\hat{p} = \frac{x}{n} = \frac{47}{100} = 0.47.$$

1.2 Question 2 — Lower end of a 95% CI for p

Question. Using the central limit theorem as an approximation, construct a 95% confidence interval for p , the probability of heads, and report the *lower* end (rounded to two decimals).

Answer.

$$\text{Lower end} \approx 0.37.$$

Explanation. Using the normal approximation for the sample proportion,

$$\hat{p} \approx N\left(p, \frac{p(1-p)}{n}\right),$$

we plug in \hat{p} for p in the standard error:

$$\hat{p} = 0.47, \quad \widehat{\text{SE}}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.47 \cdot 0.53}{100}} \approx 0.0499.$$

A 95% CI is then

$$\hat{p} \pm z_{0.975} \widehat{\text{SE}}(\hat{p}) = 0.47 \pm 1.96 \times 0.0499.$$

The lower bound is

$$0.47 - 1.96 \cdot 0.0499 \approx 0.3722 \approx 0.37.$$

1.3 Question 3 — Upper end of a 95% CI for p

Question. Using the same 95% confidence interval as in Question 2, report the *upper* end (rounded to two decimals).

Answer.

Upper end ≈ 0.57 .

Explanation. From the same calculation as above,

$$\text{upper bound} = 0.47 + 1.96 \cdot 0.0499 \approx 0.5678 \approx 0.57.$$

Thus the approximate 95% CI is

$$(0.37, 0.57).$$

1.4 Question 4 — Definition of the likelihood

Question. The likelihood function for parameter θ with data y is based on which of the following?

- $P(\theta \mid y)$
- $P(y \mid \theta)$
- $P(\theta)$
- $P(y)$
- None of the above.

Answer.

The correct choice is $P(y \mid \theta)$.

Explanation. The *likelihood function* for θ given observed data y is

$$L(\theta \mid y) = P(y \mid \theta)$$

viewed as a function of θ (with y fixed). - $P(\theta \mid y)$ is the posterior (Bayesian). - $P(\theta)$ is the prior. - $P(y)$ is the marginal probability of the data.

1.5 Question 5 — MLE for λ in an Exponential model

Question. If $X_1, \dots, X_n \stackrel{\text{iid}}{\sim} \text{Exponential}(\lambda)$, the MLE for λ is

$$\hat{\lambda} = \frac{1}{\bar{x}},$$

where \bar{x} is the sample mean. Suppose we observe

$$X_1 = 2.0, \quad X_2 = 2.5, \quad X_3 = 4.1, \quad X_4 = 1.8, \quad X_5 = 4.0.$$

Compute the MLE $\hat{\lambda}$, rounded to two decimal places.

Answer.

$$\hat{\lambda} \approx 0.35.$$

Explanation. Compute the sample mean:

$$\bar{x} = \frac{2.0 + 2.5 + 4.1 + 1.8 + 4.0}{5} = \frac{14.4}{5} = 2.88.$$

Then the MLE is

$$\hat{\lambda} = \frac{1}{\bar{x}} = \frac{1}{2.88} \approx 0.3472 \approx 0.35.$$

1.6 Question 6 — MLE for μ in a Normal model

Question. If X_1, \dots, X_n are iid from a Normal distribution $N(\mu, \sigma^2)$ with unknown mean μ , then the MLE for μ is the sample mean \bar{x} . Suppose we observe $n = 4$ data points

$$x = \{-1.2, 0.5, 0.8, -0.3\}.$$

What is the MLE $\hat{\mu}$? Round your answer to two decimal places.

Answer.

$$\hat{\mu} = -0.05.$$

Explanation. The sample mean is

$$\bar{x} = \frac{-1.2 + 0.5 + 0.8 - 0.3}{4} = \frac{-1.2 + 0.5}{4} + \frac{0.8 - 0.3}{4} = \frac{-0.7}{4} + \frac{0.5}{4} = \frac{-0.2}{4} = -0.05.$$

Hence the MLE is

$$\hat{\mu} = \bar{x} = -0.05.$$

2 Exam 2

2.1 Questions 1–5: Political Preferences

2.1.1 Question 1

Question. You ask 5 yes/no questions, each “yes” indicating a conservative viewpoint. Let Y be the number of “yes” responses and

$$\theta \in \{\text{conservative}, \text{liberal}\}.$$

If $\theta = \text{conservative}$, then $P(\text{yes}) = 0.8$. If $\theta = \text{liberal}$, then $P(\text{no}) = 0.7$ so $P(\text{yes}) = 0.3$. Assume independence across the 5 questions. What is an appropriate likelihood for this scenario?

Answer.

$$f(y | \theta) = \binom{5}{y} 0.8^y 0.2^{5-y} 1\{\theta = \text{conservative}\} + \binom{5}{y} 0.3^y 0.7^{5-y} 1\{\theta = \text{liberal}\}.$$

Explanation. Given $\theta = \text{conservative}$, $Y | \theta \sim \text{Binomial}(5, 0.8)$, hence

$$P(Y = y | \theta = \text{conservative}) = \binom{5}{y} 0.8^y 0.2^{5-y}.$$

Given $\theta = \text{liberal}$, $Y | \theta \sim \text{Binomial}(5, 0.3)$, hence

$$P(Y = y | \theta = \text{liberal}) = \binom{5}{y} 0.3^y 0.7^{5-y}.$$

We combine both in a single likelihood expression using indicator functions.

2.1.2 Question 2

Question. Your colleague answers “no” to all 5 questions, so $Y = 0$. What is the MLE for θ ?

Answer.

$$\hat{\theta} = \text{liberal}.$$

Explanation. Compute the likelihood under each value of θ :

$$P(Y = 0 | \theta = \text{conservative}) = \binom{5}{0} 0.8^0 0.2^5 = 0.2^5 = 0.00032,$$

$$P(Y = 0 \mid \theta = \text{liberal}) = \binom{5}{0} 0.3^0 0.7^5 = 0.7^5 \approx 0.16807.$$

The MLE is the value of θ that maximizes the likelihood, i.e. $\theta = \text{liberal}$.

2.1.3 Question 3

Question. With prior

$$P(\theta = \text{conservative}) = 0.5, \quad P(\theta = \text{liberal}) = 0.5,$$

Bayes' theorem gives

$$f(\theta \mid y) = \frac{f(y \mid \theta)f(\theta)}{\sum_{\theta'} f(y \mid \theta')f(\theta')}.$$

What is the corresponding expression for this problem?

Answer.

$$f(\theta \mid y) = \frac{\binom{5}{y} 0.8^y 0.2^{5-y} (0.5) 1\{\theta = \text{conservative}\} + \binom{5}{y} 0.3^y 0.7^{5-y} (0.5) 1\{\theta = \text{liberal}\}}{\binom{5}{y} 0.8^y 0.2^{5-y} (0.5) + \binom{5}{y} 0.3^y 0.7^{5-y} (0.5)}.$$

Explanation. We plug in:

$$f(y \mid \theta = \text{conservative}) = \binom{5}{y} 0.8^y 0.2^{5-y}, \quad f(y \mid \theta = \text{liberal}) = \binom{5}{y} 0.3^y 0.7^{5-y},$$

and $f(\theta) = 0.5$ for each, then normalize over the two possible values of θ .

2.1.4 Question 4

Question. Evaluate the posterior probability $P(\theta = \text{conservative} \mid Y = 0)$ and round to three decimal places.

Answer.

$$P(\theta = \text{conservative} \mid Y = 0) \approx 0.002.$$

Explanation. Using $y = 0$:

$$f(y = 0 \mid \theta = \text{conservative}) = 0.2^5 = 0.00032, \quad f(y = 0 \mid \theta = \text{liberal}) = 0.7^5 \approx 0.16807.$$

With prior $P(\theta = \text{conservative}) = P(\theta = \text{liberal}) = 0.5$,

$$\text{num}_{\text{cons}} = 0.5 \cdot 0.00032 = 0.00016, \quad \text{num}_{\text{lib}} = 0.5 \cdot 0.16807 = 0.084035.$$

Normalization constant:

$$Z = 0.00016 + 0.084035 = 0.084195.$$

Thus

$$P(\theta = \text{conservative} \mid Y = 0) = \frac{0.00016}{0.084195} \approx 0.0019 \approx 0.002.$$

2.1.5 Question 5

Question. Evaluate the posterior probability $P(\theta = \text{liberal} \mid Y = 0)$ and round to three decimal places.

Answer.

$$P(\theta = \text{liberal} \mid Y = 0) \approx 0.998.$$

Explanation. Either compute directly:

$$P(\theta = \text{liberal} \mid Y = 0) = \frac{0.084035}{0.084195} \approx 0.9981 \approx 0.998,$$

or use that the posterior probabilities must sum to 1:

$$P(\theta = \text{liberal} \mid Y = 0) = 1 - P(\theta = \text{conservative} \mid Y = 0) \approx 1 - 0.002 = 0.998.$$

2.2 Questions 6–9: Loaded Coins

2.2.1 Question 6

Question. You have three possible coins:

$$\theta \in \{\text{fair, loaded favoring heads, loaded favoring tails}\}.$$

Let X be the number of heads in 4 flips.

- Fair coin: $P(\text{heads}) = 0.5$.
- Loaded favoring heads: $P(\text{heads}) = 0.7$.
- Loaded favoring tails: $P(\text{tails}) = 0.7$, so $P(\text{heads}) = 0.3$.

What is the form of the likelihood $f(x \mid \theta)$?

Answer.

$$\begin{aligned} f(x | \theta) &= \binom{4}{x} 0.5^x 0.5^{4-x} 1\{\theta = \text{fair}\} \\ &\quad + \binom{4}{x} 0.7^x 0.3^{4-x} 1\{\theta = \text{loaded heads}\} \\ &\quad + \binom{4}{x} 0.3^x 0.7^{4-x} 1\{\theta = \text{loaded tails}\}. \end{aligned}$$

Explanation. Under each coin, X is Binomial(4, p) with different p :

$$p = \begin{cases} 0.5, & \theta = \text{fair}, \\ 0.7, & \theta = \text{loaded heads}, \\ 0.3, & \theta = \text{loaded tails}. \end{cases}$$

We write the binomial pmf for each case and combine with indicator functions.

2.2.2 Question 7

Question. Suppose you flip the coin 4 times and observe $X = 2$ heads. What is the MLE for θ ?

Answer.

$$\hat{\theta} = \text{fair}.$$

Explanation. Compute the likelihoods for $x = 2$:

$$\binom{4}{2} = 6.$$

$$L_{\text{fair}} = P(X = 2 | \theta = \text{fair}) = 6 \cdot 0.5^2 0.5^2 = 6 \cdot 0.5^4 = 6 \cdot 0.0625 = 0.375.$$

$$L_{\text{heads}} = P(X = 2 | \theta = \text{loaded heads}) = 6 \cdot 0.7^2 0.3^2 = 6 \cdot 0.49 \cdot 0.09 = 6 \cdot 0.0441 = 0.2646.$$

$$L_{\text{tails}} = P(X = 2 | \theta = \text{loaded tails}) = 6 \cdot 0.3^2 0.7^2 = 0.2646.$$

Since $L_{\text{fair}} = 0.375$ is largest, the MLE is $\theta = \text{fair}$.

2.2.3 Question 8

Question. With prior

$$P(\theta = \text{fair}) = 0.4, \quad P(\theta = \text{loaded heads}) = 0.3, \quad P(\theta = \text{loaded tails}) = 0.3,$$

and observing $X = 2$ heads in 4 flips, what is the posterior probability that the coin is fair? Round to two decimal places.

Answer.

$$P(\theta = \text{fair} \mid X = 2) \approx 0.49.$$

Explanation. We already have

$$L_{\text{fair}} = 0.375, \quad L_{\text{heads}} = 0.2646, \quad L_{\text{tails}} = 0.2646.$$

Multiply by priors:

$$\text{num}_{\text{fair}} = 0.4 \cdot 0.375 = 0.15,$$

$$\text{num}_{\text{heads}} = 0.3 \cdot 0.2646 = 0.07938,$$

$$\text{num}_{\text{tails}} = 0.3 \cdot 0.2646 = 0.07938.$$

Normalization constant:

$$Z = 0.15 + 0.07938 + 0.07938 = 0.30876.$$

Hence

$$P(\theta = \text{fair} \mid X = 2) = \frac{0.15}{0.30876} \approx 0.4858 \approx 0.49.$$

2.2.4 Question 9

Question. With the same data $X = 2$ and priors, what is the posterior probability that the coin is *loaded* (favoring either heads or tails)? Round to two decimal places.

Answer.

$$P(\theta = \text{loaded} \mid X = 2) \approx 0.51.$$

Explanation. The posterior probability that the coin is loaded (either heads or tails) is

$$P(\theta = \text{loaded} \mid X = 2) = P(\theta = \text{loaded heads} \mid X = 2) + P(\theta = \text{loaded tails} \mid X = 2).$$

By complement,

$$P(\theta = \text{loaded} \mid X = 2) = 1 - P(\theta = \text{fair} \mid X = 2) \approx 1 - 0.4858 \approx 0.5142 \approx 0.51.$$

Equivalently, one could compute directly:

$$\frac{0.07938 + 0.07938}{0.30876} \approx 0.5142 \approx 0.51.$$

3 Continuous Bayes, Credible Intervals

3.1 Question 1 — When to use the continuous Bayes formula

Question. We use the continuous version of Bayes' theorem if:

- θ is continuous
- Y is continuous
- $f(y | \theta)$ is continuous
- All of the above
- None of the above

Answer.

We use the continuous version if θ is continuous.

Explanation. The continuous Bayes' theorem has the form

$$f(\theta | y) = \frac{f(y | \theta)f(\theta)}{\int f(y | \theta)f(\theta) d\theta},$$

which involves an *integral over* θ . This is the continuous case in the *parameter* θ . It does not require Y to be continuous; Y may be discrete (e.g. Bernoulli) with a continuous prior on θ .

3.2 Question 2 — Sequential update after two heads

Question. Coin-flip example:

- Likelihood for one flip: $Y \in \{0, 1\}$,

$$f(y | \theta) = \theta^y (1 - \theta)^{1-y} 1_{\{0 \leq \theta \leq 1\}},$$

- Prior: Uniform(0, 1). - After the first flip $Y_1 = 1$, the posterior is

$$f(\theta | Y_1 = 1) = 2\theta 1_{\{0 \leq \theta \leq 1\}}.$$

Now use this posterior as the prior for θ before the second flip. The second flip also results in heads, $Y_2 = 1$. Which expression represents the posterior PDF after the second head?

Given options include (most relevant one):

$$f(\theta \mid Y_2 = 1) = \frac{\theta \cdot 2\theta}{\int_0^1 \theta \cdot 2\theta d\theta} 1_{\{0 \leq \theta \leq 1\}}.$$

Answer.

$$f(\theta \mid Y_1 = 1, Y_2 = 1) = \frac{\theta \cdot 2\theta}{\int_0^1 \theta \cdot 2\theta d\theta} 1_{\{0 \leq \theta \leq 1\}}.$$

Explanation. Sequential Bayes update:

$$\text{new posterior} \propto \text{likelihood of new data} \times \text{current prior}.$$

Here, for the second flip $Y_2 = 1$:

$$f(y_2 = 1 \mid \theta) = \theta, \quad \text{and prior density } \pi(\theta) = 2\theta.$$

So the unnormalized posterior is

$$\theta \cdot 2\theta = 2\theta^2, \quad 0 \leq \theta \leq 1.$$

Normalize:

$$\int_0^1 2\theta^2 d\theta = 2 \left[\frac{\theta^3}{3} \right]_0^1 = \frac{2}{3}, \quad \Rightarrow \quad f(\theta \mid Y_1 = 1, Y_2 = 1) = 3\theta^2, \quad 0 \leq \theta \leq 1,$$

which is exactly

$$\frac{\theta \cdot 2\theta}{\int_0^1 \theta \cdot 2\theta d\theta} 1_{\{0 \leq \theta \leq 1\}}.$$

3.3 Question 3 — Interpretation of a prior probability

Question. We use a $\text{Uniform}(0, 1)$ prior for θ and know that

$$P(0.3 < \theta < 0.9) = 0.6.$$

Which is the correct interpretation?

Answer.

$(0.3, 0.9)$ is a 60% credible interval for θ *before* observing any data.

Explanation. Since this probability comes purely from the *prior* ($\text{Uniform}(0, 1)$), it is a statement about prior uncertainty:

$$P_{\text{prior}}(0.3 < \theta < 0.9) = 0.6.$$

It is not a confidence interval (frequentist) and not based on any observed data.

3.4 Question 4 — Interpretation of a posterior probability

Question. After observing $Y = 1$, the posterior is

$$f(\theta \mid Y = 1) = 2\theta \mathbf{1}_{\{0 \leq \theta \leq 1\}}.$$

We compute

$$P(0.3 < \theta < 0.9 \mid Y = 1) = \int_{0.3}^{0.9} 2\theta \, d\theta = [\theta^2]_{0.3}^{0.9} = 0.81 - 0.09 = 0.72.$$

Which is the correct interpretation?

Answer.

$(0.3, 0.9)$ is a 72% credible interval for θ *after* observing $Y = 1$.

Explanation. This is a *posterior* probability statement:

$$P(\theta \in (0.3, 0.9) \mid Y = 1) = 0.72.$$

So $(0.3, 0.9)$ is a credible interval with 72% posterior probability. It is not a frequentist confidence interval and not a prior probability.

3.5 Question 5 — Quantiles for a 90% equal-tailed interval

Question. Which two quantiles capture the middle 90% of a distribution in an *equal-tailed* sense?

Answer.

0.05 and 0.95.

Explanation. A 90% equal-tailed interval leaves 5% in each tail:

$$P(\theta \leq q_{0.05}) = 0.05, \quad P(\theta \geq q_{0.95}) = 0.05,$$

so the middle 90% is between the 0.05 and 0.95 quantiles.

3.6 Question 6 — 95% equal-tailed interval for $N(0, 1)$

Question. Suppose your posterior is

$$\theta \mid y \sim N(0, 1).$$

Report the *upper* end of a 95% equal-tailed interval for θ (rounded to two decimals).

Answer.

Upper end ≈ 1.96 .

Explanation. For a standard Normal, the 95% equal-tailed interval is

$$[\Phi^{-1}(0.025), \Phi^{-1}(0.975)] \approx [-1.96, 1.96].$$

Thus the upper bound is 1.96 to two decimal places.

3.7 Question 7 — Meaning of HPD

Question. What does “HPD interval” stand for?

Answer.

HPD interval = Highest Posterior Density interval.

Explanation. An HPD interval is the set (often an interval in 1D) with posterior probability $1 - \alpha$ such that every point inside has *higher* posterior density than any point outside. It is the Bayesian analogue of “take the top $(1 - \alpha)$ mass of the posterior.”

3.8 Question 8 — 50% credible interval

Question. Compute the 50% credible interval from a posterior distribution.

General Answer (equal-tailed). Let F be the posterior CDF of θ and let $q_p = F^{-1}(p)$ be the p -th posterior quantile. A 50% equal-tailed credible interval is

$$[q_{0.25}, q_{0.75}],$$

i.e., the interval between the 25% and 75% posterior quantiles.

Example (if $\theta \mid y \sim N(0, 1)$). For a standard Normal posterior,

$$q_{0.25} \approx -0.67, \quad q_{0.75} \approx 0.67,$$

so a 50% equal-tailed credible interval is

$$[-0.67, 0.67].$$

4 Bayesian vs frequentist

4.1 Basic Interpretations

Frequentist confidence interval. A $(1 - \alpha)$ confidence interval procedure $C_f(Y)$ satisfies

$$P_\theta(\theta \in C_f(Y)) \approx 1 - \alpha \quad \text{for all relevant } \theta,$$

where the probability is over repeated samples Y drawn from the model with *fixed* true parameter θ .

Interpretation:

If we repeated the data collection and interval computation many times, about $100(1 - \alpha)\%$ of the resulting intervals would contain the true parameter; for the one observed interval, the true value is either inside or outside, but we do not assign a probability to that event.

Bayesian credible interval. Let $\pi(\theta)$ be a prior and $\pi(\theta | y)$ the posterior. A $(1 - \alpha)$ credible set $C_B(y)$ satisfies

$$P(\theta \in C_B(y) | y) = \int_{C_B(y)} \pi(\theta | y) d\theta = 1 - \alpha.$$

Interpretation:

Given the data y and the prior model, your posterior probability that θ lies in $C_B(y)$ is $100(1 - \alpha)\%$.

4.2 When to Prefer Frequentist Confidence Intervals

4.2.1 (a) Regulatory or standardized settings

In drug trials, industrial quality control, or other regulated environments, requirements are often specified in frequentist terms:

“Use a 95% confidence interval,” i.e., a procedure with 95% coverage.

If the goal is to guarantee long-run error control across repeated studies, coverage is the natural criterion, so frequentist intervals are preferred.

4.2.2 (b) Need for guaranteed coverage properties

If your main design question is

“Which rule C_f has $P_\theta(\theta \in C_f(Y)) \geq 1 - \alpha$ for all θ ?”

then you are explicitly optimizing a frequentist property: robustness over all parameter values without relying on a specific prior.

4.2.3 (c) Avoiding or lacking a prior

In contentious settings where analysts cannot agree on a prior, or where you wish to be “prior-free”:

- Frequentist intervals provide a neutral baseline.
- Everyone can agree on their coverage interpretation even if they disagree about prior beliefs.

4.2.4 (d) Sample size and design calculations

Power and sample-size planning are traditionally frequentist:

$$P_\theta(\text{reject } H_0) = \text{desired power},$$

often connected to the width or behavior of confidence intervals. In such contexts, using frequentist intervals is natural and consistent with the rest of the design.

4.3 When to Prefer Bayesian Credible Intervals

4.3.1 (a) Direct probability statements about parameters

If you want to answer questions like

$$P(\theta > 0 \mid y) \quad \text{or} \quad P(a < \theta < b \mid y) = 0.95,$$

then Bayesian credible intervals are the appropriate tool. They directly encode your uncertainty about θ given the data and prior.

4.3.2 (b) Presence of genuine prior information

When substantial prior information exists (historical data, physical constraints, expert knowledge), Bayesian methods allow you to incorporate this via $\pi(\theta)$ and obtain posterior intervals that com-

bine prior and data:

$$\pi(\theta | y) \propto f(y | \theta) \pi(\theta).$$

In small samples, this can stabilize inference and produce more reasonable intervals than purely likelihood-based ones.

4.3.3 (c) Complex or hierarchical models

For multilevel/hierarchical models, random effects, or high-dimensional latent structures, frequentist confidence intervals can be hard to derive or approximate. Bayesian posterior computation (e.g. MCMC) naturally yields marginal posterior distributions for many parameters, from which credible intervals are straightforward to obtain.

4.3.4 (d) Decision-making with losses or utilities

In decision-theoretic problems with a loss function $L(\theta, a)$, actions are chosen based on posterior expectations:

$$\mathbb{E}[L(\theta, a) | y] = \int L(\theta, a) \pi(\theta | y) d\theta.$$

Here, probabilities like $P(\theta \in C_B(y) | y)$ play a direct role in choosing optimal actions, so Bayesian credible intervals align with the decision framework.

4.4 When They Coincide (Approximately)

In many regular problems with large sample size:

- Likelihood is approximately Normal in θ ,
- Priors are weak or diffuse,

then

$$\pi(\theta | y) \approx N(\hat{\theta}, \widehat{\text{Var}}(\hat{\theta})),$$

and the $(1 - \alpha)$ Bayesian credible interval often numerically matches the $(1 - \alpha)$ frequentist confidence interval. In such cases, the difference is mostly interpretational rather than numerical.

4.5 Compact Summary

Let

- $C_f(Y)$ be a frequentist confidence interval rule,
- $C_B(Y)$ be a Bayesian credible set rule.

Then:

Prefer frequentist CIs $C_f(Y)$	when you prioritize long-run coverage	$(P_\theta(\theta \in C_f(Y)) \approx 1 - \alpha)$.
Prefer Bayesian credible sets $C_B(Y)$	when you prioritize posterior belief	$(P(\theta \in C_B(Y) \mid Y = y) = 1 - \alpha)$.

5 Honor Quizz

5.1 Question 1

Question. Although the likelihood function is not always a product of

$$L(\theta \mid y_1, \dots, y_n) = \prod_{i=1}^n f(y_i \mid \theta),$$

this product form is convenient mathematically. What assumption about the observations y_1, \dots, y_n allows us to multiply their individual likelihood components?

Answer.

The observations are independent (given θ).

Explanation. By definition, if Y_1, \dots, Y_n are independent given θ , then

$$P(Y_1 = y_1, \dots, Y_n = y_n \mid \theta) = \prod_{i=1}^n P(Y_i = y_i \mid \theta) = \prod_{i=1}^n f(y_i \mid \theta).$$

So independence (conditional on θ) justifies writing the joint likelihood as a product of marginal likelihoods.

5.2 Question 2

Question. One property of maximum likelihood estimators (MLEs) is transformation invariance: if $\hat{\theta}$ is the MLE for θ , then the MLE for $g(\theta)$ is $g(\hat{\theta})$ for any function $g(\cdot)$.

Suppose you conduct $n = 25$ Bernoulli trials and observe $x = 10$ successes. What is the MLE for the *odds* of success

$$g(\theta) = \frac{\theta}{1 - \theta}?$$

Answer.

$$\widehat{\text{odds}} = \frac{2}{3} \approx 0.67.$$

Explanation. The Bernoulli success probability θ has likelihood

$$L(\theta \mid x) \propto \theta^x (1 - \theta)^{n-x},$$

and the MLE is the sample proportion:

$$\hat{\theta} = \frac{x}{n} = \frac{10}{25} = 0.4.$$

By transformation invariance of MLEs, the MLE of the odds $g(\theta) = \theta/(1 - \theta)$ is

$$\widehat{g(\theta)} = g(\hat{\theta}) = \frac{\hat{\theta}}{1 - \hat{\theta}} = \frac{0.4}{1 - 0.4} = \frac{0.4}{0.6} = \frac{2}{3} \approx 0.67.$$

5.3 Question 3

Question. A coin may be fair (probability of heads 0.5) or loaded (probability of heads 0.7). Let L denote “loaded” and F denote “fair”.

Your prior probabilities are

$$P(L) = 0.6, \quad P(F) = 0.4.$$

Another sibling proposes a bet:

- If the coin is loaded (L), she pays you \$1.
- If the coin is fair (F), you pay her \$ z .

Find z that makes the game fair in terms of your prior (i.e., prior expected payoff = 0).

Answer.

$$z = 1.5.$$

Explanation. Let X be your payoff. Then

$$X = \begin{cases} +1, & \text{if } L, \\ -z, & \text{if } F. \end{cases}$$

The prior expected payoff is

$$\mathbb{E}[X] = P(L) \cdot 1 + P(F) \cdot (-z) = 0.6 - 0.4z.$$

A fair game (in your prior sense) requires $\mathbb{E}[X] = 0$:

$$0.6 - 0.4z = 0 \implies 0.4z = 0.6 \implies z = \frac{0.6}{0.4} = 1.5.$$

5.4 Question 4

Question. Before taking the bet, you agree to flip the coin once. It lands heads (H).

Your sister now argues this is evidence for the loaded coin and demands to increase z to 2:

- If L (loaded): you gain \$1.
- If F (fair): you lose \$2.

Using your *posterior* probability after observing one head, should you accept the new bet at $z = 2$?

Answer (correct option).

Yes, your posterior expected payoff is now greater than \$0.

Explanation. First update your belief using Bayes' rule after observing one head:

$$\begin{aligned} P(H | L) &= 0.7, & P(H | F) &= 0.5, \\ P(L) &= 0.6, & P(F) &= 0.4. \end{aligned}$$

Joint (unnormalized) probabilities:

$$P(L, H) = P(H | L)P(L) = 0.7 \cdot 0.6 = 0.42,$$

$$P(F, H) = P(H | F)P(F) = 0.5 \cdot 0.4 = 0.20.$$

Marginal probability of heads:

$$P(H) = P(L, H) + P(F, H) = 0.42 + 0.20 = 0.62.$$

Posterior probabilities:

$$P(L | H) = \frac{P(L, H)}{P(H)} = \frac{0.42}{0.62} \approx 0.6774,$$

$$P(F | H) = \frac{P(F, H)}{P(H)} = \frac{0.20}{0.62} \approx 0.3226.$$

With $z = 2$, your payoff X is

$$X = \begin{cases} +1, & \text{if } L, \\ -2, & \text{if } F. \end{cases}$$

Posterior expected payoff:

$$E[X \mid H] = P(L \mid H) \cdot 1 + P(F \mid H) \cdot (-2) \approx 0.6774 - 2(0.3226) \approx 0.6774 - 0.6452 \approx 0.0322 > 0.$$

Since your posterior expected payoff is positive, you *should* accept the new bet at $z = 2$ according to your Bayesian posterior.