Frequentist and Bayesian Inference: Practice

Section 10 handout

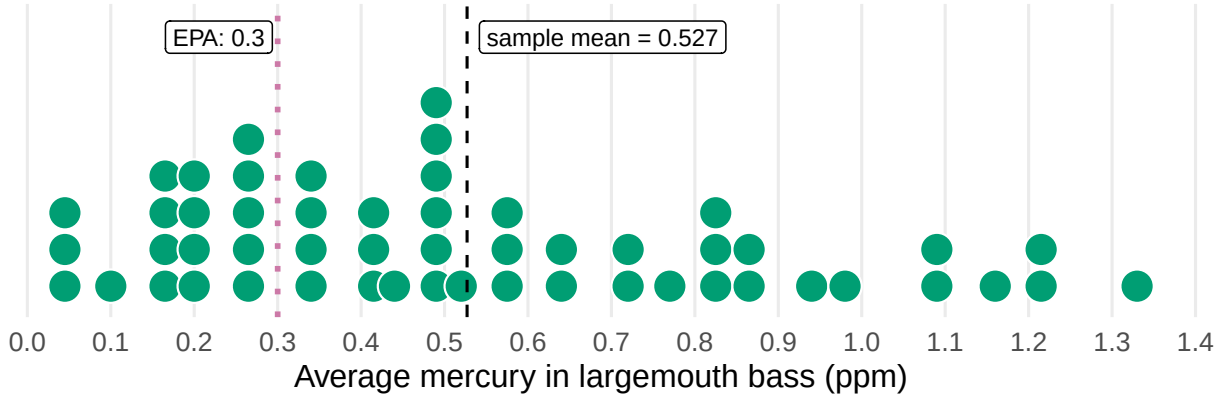
Name: _____

Is the bass safe to eat?

Mercury builds up in fish, and people who eat a lot of fish with high mercury levels can be harmed. The U.S. Environmental Protection Agency (EPA) set **0.3 ppm** (parts per million, or mg of methylmercury per kg of fish) as the level of mercury in fish tissue that protects the health of people who eat fish.

Researchers caught **largemouth bass** from 53 lakes in Florida and measured the mercury in each fish. For each lake, we have the **average mercury level** (in ppm) of the bass from that lake.

Data: `FloridaLakes` in the `Lock5Data` package. Lange, T. R., Royals, H. E., & Connor, L. L. (1993). Influence of water chemistry on mercury concentration in largemouth bass from Florida lakes. *Transactions of the American Fisheries Society*, 122(1), 74–84. U.S. EPA (2001). *Water Quality Criterion for the Protection of Human Health: Methylmercury* (EPA-823-R-01-001).



Sample mean: $\bar{x} = 0.527$ ppm. Standard error: 0.047 (given to you).

Let μ be the mean of the average mercury levels in largemouth bass across **all** Florida lakes (like the ones in the study).

Is $\bar{x}$ a statistic or a parameter? _____

Is μ a statistic or a parameter? _____

Today we will analyze the same data in two ways: first the **frequentist** way (Week 9), then the **Bayesian** way (Week 10). Pay attention to **what is random** and **what each probability is about**.

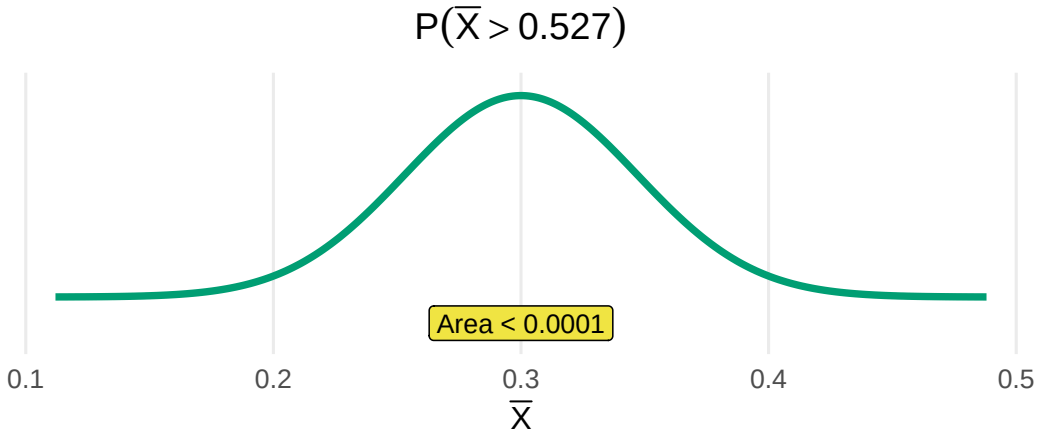
Part 1: The frequentist analysis

Hypotheses.

- H_0: the mean mercury level is at the EPA level. $\mu =$ _____
- H_A: the mean mercury level is not at the EPA level. μ _____

In frequentist inference, μ is a _____, and we start by _____.

Null sampling distribution. If μ were 0.3, the sample mean of 53 lakes would be approximately $\bar{X} \mid \mu = 0.3 \sim N(0.3, 0.047^2)$. The plot shows this distribution with the area above 0.527 shaded.



a. Where is $\bar{x} = 0.527$ on this plot? About how many standard errors is it above 0.3?

b. Write the p-value in conditional probability notation. Remember to count both tails.

c. Use the plot to find the p-value. _____

d. What is your conclusion at $\alpha = 0.05$? Write it in context.

e. Which type of error could you have made? Describe it in context.

Confidence interval. The margin of error for a 95% confidence interval is 0.092.

f. Write the 95% confidence interval. _____

g. Interpret the interval in context.

h. Is 0.3 inside the interval? How does this agree with your conclusion in (d)?

i. What does "95%" in "95% confidence" describe: this one interval, or the method? Explain.

j. Is the result practically significant? Does it mean that the bass in **every** Florida lake is unsafe to eat? Explain.

Part 2: The Bayesian analysis

Jordan, a state official, analyzes the same data the Bayesian way. Before seeing the data, Jordan believes that the bass are probably close to the EPA level. Jordan's _____ is

$$\mu \sim N(0.3, 0.1^2)$$

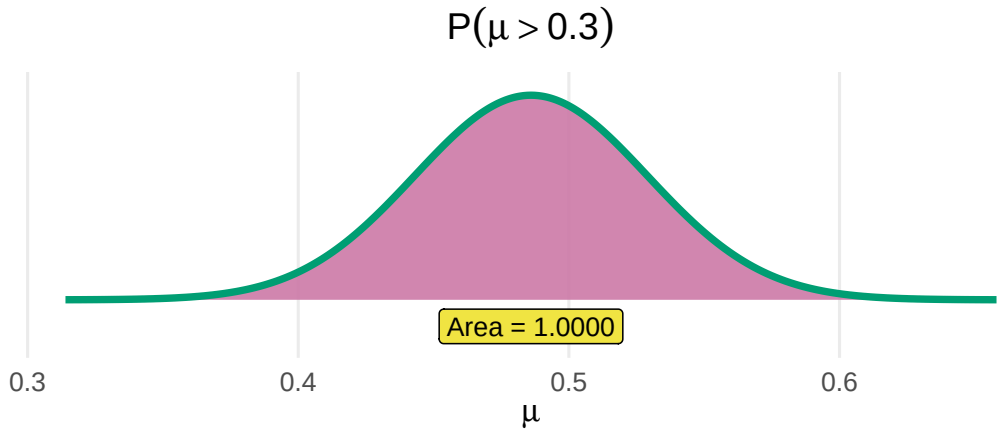
In Bayesian inference, μ is a _____: its distribution describes what we believe about it.

Hypotheses. The Bayesian question: is the mean mercury level **above** the EPA level?

- H_0: $\mu \leq 0.3$. The mean mercury level is at or below the EPA level.
- H_A: $\mu > 0.3$. The mean mercury level is above the EPA level.

a. What is Jordan's **prior** probability of H_A, $P(\mu > 0.3)$? How do you know without a calculation?

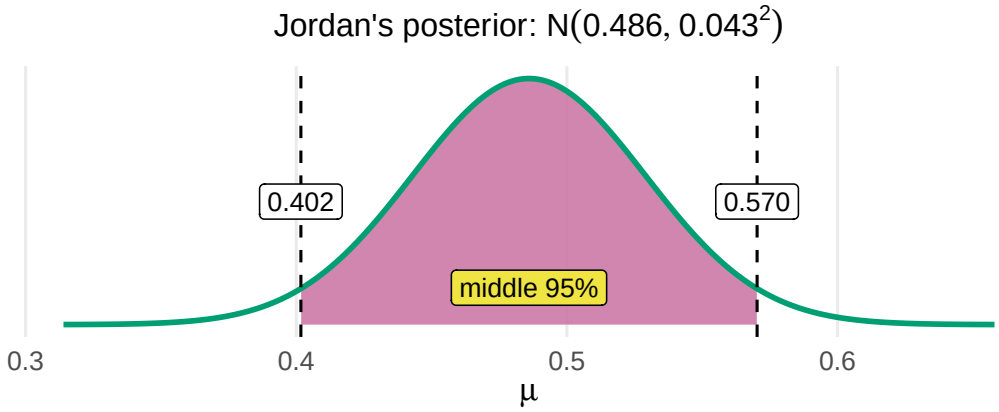
After seeing the 53 lakes, Jordan's _____ is $\mu \mid \text{data} \sim N(0.486, 0.043^2)$. The plot shows this distribution with the area above 0.3 shaded.



b. Fill in: $P(H_A \mid \text{data}) = P(\mu > 0.3 \mid \text{data})$ _____ and $P(H_0 \mid \text{data})$ _____

c. Write a conclusion in context.

Credible interval. The plot shows Jordan's posterior with the middle 95% shaded.



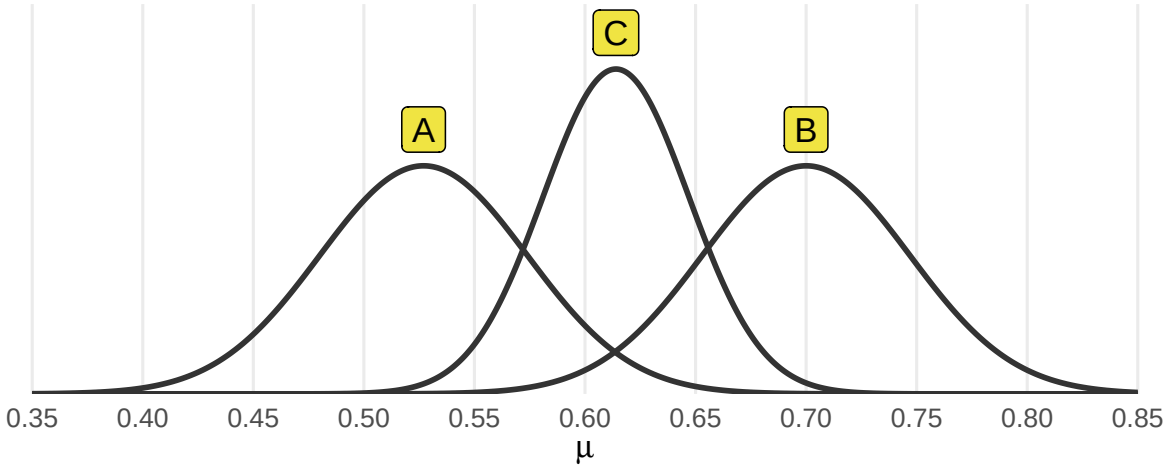
d. Jordan's 95% credible interval is from _____ to _____ ppm.

e. Interpret the interval in context, and write it in probability notation.

f. Jordan's posterior mean, 0.486, is between the prior mean, 0.3, and the sample mean, 0.527. Which one is it closer to? Why?

A different prior. Riley, an environmental scientist, expects mercury in Florida bass to be high. Riley's prior is $\mu \sim N(0.7, 0.047^2)$. Riley sees the same 53 lakes.

The plot shows Riley's prior, the likelihood, and Riley's posterior, labeled A, B, and C (not in that order).



g. Which curve is Riley's prior? _____ The likelihood? _____ Riley's posterior? _____

h. Explain how you decided.

i. Riley's prior (standard deviation 0.047) and the likelihood (standard error 0.047) are equally certain. Riley's posterior mean is the weighted average _____.

j. Jordan's posterior mean is 0.486 and Riley's is 0.614. They saw the same data. Why are their posteriors different? What would bring them closer together?

Part 3: Same data, two kinds of answers

Which inference? For each statement about the mercury data, write **F** (frequentist) or **B** (Bayesian), and write the probability in notation.

Table 1: Frequentist or Bayesian? Statements about the mercury data

Statement	F or B?	In notation
If μ were 0.3, there would be less than a 0.1% chance of a sample mean as far from 0.3 as 0.527.	_____	_____
Given the data, it is almost certain (probability more than 0.9999) that the mean mercury level is above 0.3 ppm.	_____	_____
Before seeing the data, Jordan thought there was a 50% chance that the mean mercury level is above 0.3 ppm.	_____	_____
If we repeated the study many times, about 95% of the intervals built this way would capture μ.	_____	_____
There is a 95% probability that μ is between 0.402 and 0.570 ppm.	_____	_____
If μ were 0.3, there would be a 5% chance that we wrongly reject H_0.	_____	_____

Look at your notation column. What do all of the frequentist probabilities have in common? What about the Bayesian ones?

Same curve, different meaning. Both of these are normal curves that we used today. Complete the table.

Table 2: Comparing the null sampling distribution and the posterior distribution

	$\bar{X} \mid \mu = 0.3 \sim N(0.3, 0.047^2)$	$\mu \mid \text{data} \sim N(0.486, 0.043^2)$
Name of this distribution	_____	_____
What is on the x-axis (what is random)?	_____	_____
What is fixed (after the |)?	_____	_____
Frequentist or Bayesian?	_____	_____
An area under it is a probability about	_____	_____

Two tiny numbers. The p-value is less than 0.001, and Jordan's $P(H_0 \mid \text{data})$ is less than 0.0001. A classmate says: "Both are tiny, so they mean the same thing: there is almost no chance that H_0 is true." Do you agree? Explain.

Part 4: Spot the error

Each statement about the mercury data is **not correct**. Explain what is wrong, and fix it.

a. "The p-value is less than 0.001, so there is less than a 0.1% chance that the mean mercury level is 0.3 ppm."

b. "We rejected H_0, so we have proven that the mean mercury level is not 0.3 ppm."

c. "There is a 95% probability that μ is between 0.435 and 0.619 ppm." (This is the frequentist confidence interval.)

d. "95% of Florida lakes have bass with an average mercury level between 0.435 and 0.619 ppm."

e. "In Bayesian hypothesis testing, Jordan started by assuming $H_0 : \mu \leq 0.3$ was true."

f. "The likelihood is highest at $\mu = 0.527$, so there is a 100% probability that $\mu = 0.527$."

Part 5: Multiple choice

Write the letter of the best answer. Then explain your choice in one sentence.

1. Which of these is the p-value for the mercury test in Part 1?

(A) $P(\mu = 0.3 \mid \text{data})$

(B) $P(\mu > 0.3 \mid \text{data})$

(C) $P(\bar{X} \geq 0.527 \text{ or } \bar{X} \leq 0.073 \mid \mu = 0.3)$

(D) $P(\bar{X} = 0.527)$

Answer: _____

2. In frequentist inference, what is random?

(A) The parameter μ

(B) The sample mean $\bar{X}$, and intervals built from samples

(C) Both μ and $\bar{X}$

(D) Neither

Answer: _____

3. Jordan's 95% **credible** interval for μ is (0.402, 0.570). Which interpretation is correct?

(A) 95% of all samples of lakes would give an interval that captures μ.

(B) 95% of Florida lakes have bass with an average mercury level between 0.402 and 0.570 ppm.

(C) There is a 95% probability that the sample mean is between 0.402 and 0.570.

(D) Given the data and Jordan's prior, there is a 95% probability that μ is between 0.402 and 0.570.

Answer: _____

4. Jordan also asks whether μ is above 0.5 ppm. From Jordan's posterior, $P(\mu > 0.5 \mid \text{data}) \approx 0.37$ and $P(\mu \leq 0.5 \mid \text{data}) \approx 0.63$. Which is correct?

(A) Given the data and Jordan's prior, $\mu \leq 0.5$ is about 1.7 times as probable as $\mu > 0.5$: weak evidence either way.

(B) Jordan fails to reject $\mu \leq 0.5$ because $0.37 > 0.05$.

(C) If μ were 0.5, there would be a 37% chance of data this extreme.

(D) The data prove that $\mu \leq 0.5$, because $0.63 > 0.5$.

Answer: _____

5. Two researchers with very different priors analyze mercury data from a very large number of lakes. Their posteriors will most likely be

(A) still very different, because their priors were very different.

(B) almost the same.

(C) the same as their priors.

(D) impossible to find, because Bayesians must agree on a prior.

Answer: _____

6. A significance level of $\alpha = 0.05$ means

(A) $P(H_0 \text{ is true} \mid \text{reject } H_0) = 0.05$

(B) $P(H_A \text{ is true}) = 0.95$

(C) $P(\text{reject } H_0 \mid H_0 \text{ is true}) = 0.05$

(D) $P(\text{Type II error}) = 0.05$

Answer: _____

7. With the same 53 lakes ($\bar{x} = 0.527$, standard error 0.047), a frequentist test of $H_0 : \mu = 0.5$ vs. $H_A : \mu \neq 0.5$ gives a p-value of about 0.57. Which is correct?

(A) The data prove that the mean mercury level is 0.5 ppm.

(B) If μ were 0.5, results like these would not be surprising, so there is not convincing evidence that μ is different from 0.5.

(C) There is a 57% probability that $\mu = 0.5$.

(D) There is a 43% probability that $\mu \neq 0.5$.

Answer: _____

8. The first 5 lakes in the data have average mercury levels 1.23, 1.33, 0.04, 0.44, 1.20 ppm. Mark each one A if it is above 0.3 and N if not: _____. Let π be the proportion of **all** Florida lakes whose bass average above 0.3 ppm. Which value of π has the highest likelihood?

(A) 0.2

(B) 0.5

(C) 0.6

(D) 0.8

Answer: _____

$L(\pi \mid \text{data}) =$ _____. $L(0.5 \mid \text{data}) \approx$ _____ and $L(0.8 \mid \text{data}) \approx$ _____.
(Hint: $0.8^4 \approx 0.41$.)

9. In our course notation, $\mu \sim N(0.3, 0.1^2)$, with nothing after a |, most likely describes

(A) a prior distribution.

(B) a null sampling distribution.

(C) a posterior distribution.

(D) the sampling distribution of a sample mean.

Answer: _____

Exit ticket

Explain to a friend who missed class the difference between the frequentist 95% **confidence** interval (0.435, 0.619) and Jordan's 95% **credible** interval (0.402, 0.570) for the mean mercury level. Use the words **random** and **fixed**. Why is Jordan's interval shifted toward 0.3?