Questions Module 4.1

You conduct an experiment to determine how long it takes, on average, to complete a particular task. You complete the task 10 times under typical conditions and calculate the average time it takes. This experiment is an example of which type of inferential method?

- *a.* interval estimation
- *b.* hypothesis testing

Incorrect.
The correct answer is **a**.

You are estimating an unknown value, the average time to complete a task.
The results of an election poll are reported in the newspaper. The paper reports that 44% of likely voters are in favor of Candidate A, and 42% are in favor of Candidate B. The margin of error for each candidate is +/- 2%, and the confidence level is 95%. Which of the following statements is true?

- *a.* Candidate A will win the election by a margin of 2%.
- *b.* The probability that 42% of the population votes for Candidate B is 95%.
- *c.* In the election, 44% of likely voters will vote for Candidate A, and 42% will vote for Candidate B.
- *d.* With 95% confidence, the true proportion of likely voters who support Candidate A is between 0.42 and 0.46.

Incorrect.
The correct answer is **d**. This poll is the result of a sample from the larger population. You don't know the true proportion of who will vote for each candidate. You estimate the proportions using the sample.

What was the goal of Michelson's experiment?

- *a.* Calculate the average speed of light for data collected in his experiment.
- *b.* Estimate the true speed of light based on experimental results.
- *c.* Develop a better apparatus for measuring the velocity of light than Foucault.
- *d.* Prove that the speed of light is less than 300,000 km/second.

Incorrect.
The correct answer is **b**. Michelson was fundamentally interested in estimating the true speed of light using data collected during his experiment.

Interval estimation is all about inferring a population parameter using sample data. Which of the following refers to a population rather than a sample?

- *a.* noontime temperatures recorded each day for one month
- *b.* the number of milliliters of fluid dispensed by a type of bottling machine
- *c.* attendance at a single World Cup match

- *d.* the age of students in one classroom at a large university

---

Incorrect.
The correct answer is **b**.

It refers to infinite measurements from a machine (in theory, the machine will continue to dispense fluid indefinitely), rather than observed measurements taken from such a machine. The other choices all refer to subsets of a larger population or an ongoing process.

Of the following statements, which best describes the general idea of a confidence interval?

- *a.* the range of values that will contain the sample data
- *b.* the range of possible values for the population data
- *c.* the range of values likely to contain the population parameter
- *d.* the range of possible values for the sample statistic

Incorrect.
The correct answer is **c**. A confidence interval is an interval estimate for the population parameter.

The 95% confidence interval for the mean is [3, 6]. What is the correct interpretation of this interval?

- *a.* The probability that the interval [3, 6] includes the true mean is 95%.
- *b.* You are 95% confident that the true mean is in the interval [3, 6].
- *c.* The chance that the true mean falls in the interval [3, 6] is 95%.
- *d.* You are 95% certain that the sample mean falls in the interval [3, 6].

Correct.
The term "confidence" refers to the statistical confidence that the calculated interval contains the true population parameter value. The interval either includes the true parameter or it doesn't. There is no probability or "chance" associated with confidence intervals. A confidence interval is **not** the probability (or the chance) that the true parameter falls in the interval.

You collect data on 50 lots and construct the following prediction interval. What is the correct interpretation of this interval?

Prediction Interval

Parameter	Future N	Lower PI	Upper PI	1-Alpha
Individual	1	37.75066	45.02001	0.950
Mean	1	37.75066	45.02001	0.950
Std Dev	1	.	.	0.950

- *a.* The probability that the next value falls between 37.75 and 45.02 is 0.95.
- *b.* You are 95% certain that the next value will fall between 37.75 and 45.02.

- *c.* You are confident that 95% of your data fall between 37.75 and 45.02.
- *d.* You are 95% certain that the true mean is between 37.75 and 45.02.

Correct.
For a prediction interval, you are estimating the range of values for the next N samples, with a specified confidence level. In this case, N = 1. Note that this prediction interval is meaningful only if the process is stable.

You have the following tolerance interval. What is the correct interpretation of this interval?

Tolerance Intervals

Proportion	Lower PI	Upper PI	1-Alpha
0.900	37.90827	44.86239	0.950

- *a.* You are 95% certain that the probability is 0.90 that your mean is between 37.9 and 44.9.
- *b.* You are 90% certain that 95% of your sample values will fall between 37.9 and 44.9.
- *c.* You are 95% certain that at least 90% of the population values will fall between 37.9 and 44.9.
- *d.* You are 90% certain that at least 95% of the population values will fall between 37.9 and 44.9.

Incorrect.
The correct answer is **c**. For a tolerance interval, you are estimating the range of values for a proportion of your future population values, with a specified confidence level.

You test the longevity of 100 batteries. You run the batteries until they are completely depleted and record the duration for each battery. Match the question to the type of interval estimate you should use.

☐ How long will the next battery last?	a. Confidence Interval
☐ We want to offer a money back guarantee for batteries that expire before a certain number of hours. What number of hours should we guarantee?	b. Prediction Interval
☐ On average, how long do these batteries last?	c. Tolerance Interval

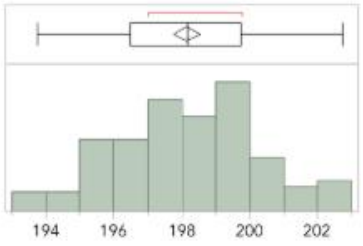
---

Incorrect.
The correct answers from top to bottom are **b**, **c**, **a**. Prediction Interval: We use sample data to estimate a single observation. Tolerance Interval: This would be a one-sided tolerance interval. We want to identify a lower bound for battery duration, where most of the batteries live longer than this value. Confidence Interval: We want an estimate of the population mean.

Consider the White Polymer case study. You might remember that one critical metric was the **Melt Flow Index (MFI)**, which has lower and upper specification limits of 192 and 198 (with a target of 195).

The data are in the file **VSSTeamData.jmp**. Using the analysis below, what conclusions can you draw about the ability to meet the target and the specification for the process?

MFI



Quantiles		
100.0%	maximum	202.74
99.5%		202.74
97.5%		202.294
90.0%		200.944
75.0%	quartile	199.725
50.0%	median	198.17
25.0%	quartile	196.48
10.0%		195.188
2.5%		193.79
0.5%		193.75
0.0%	minimum	193.75

Summary Statistics	
Mean	198.14762
Std Dev	2.1132528
Std Err Mean	0.2062324
Upper 95% Mean	198.55659
Lower 95% Mean	197.73865
N	105

Tolerance Intervals			
Proportion	Lower TI	Upper TI	1-Alpha
0.900	194.1997	202.0956	0.950

The 95% confidence interval for the mean is 197.7 to 198.6. The target is not in this interval, so the process, on average, is off target. It is shifted high. The 95%-90% tolerance interval is 194.2 to 202.1. This interval is wider than the spec limits. Unless you improve the process, you'd expect some MFI measurements to fall beyond each of the spec limits. Based on the confidence interval and the tolerance interval, you know that the process is both off target and too variable relative to the specification limits.