A Parameter in Statistics:

Concepts, Purpose, and Examples

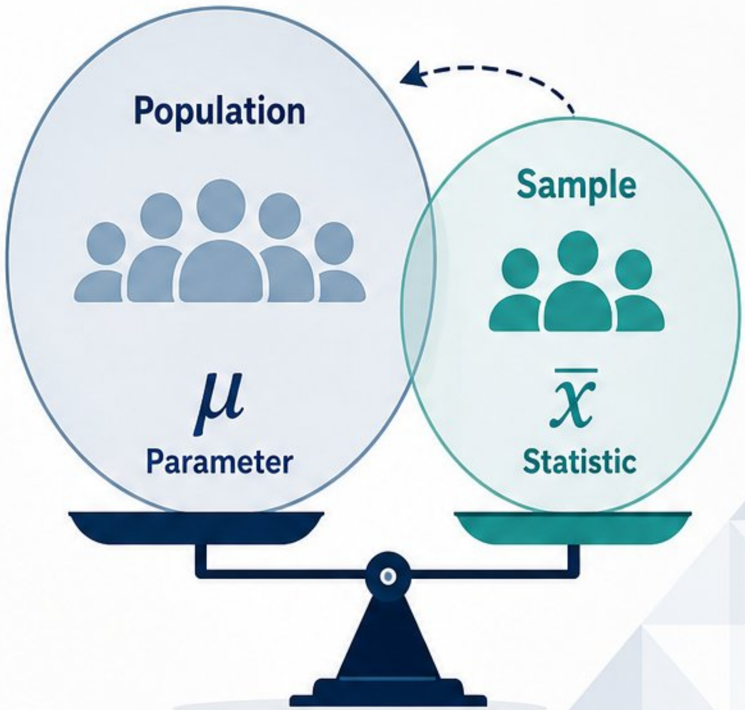
Opening hook: a sample average reported as the true population value

Objectives: separate population parameter from sample statistic

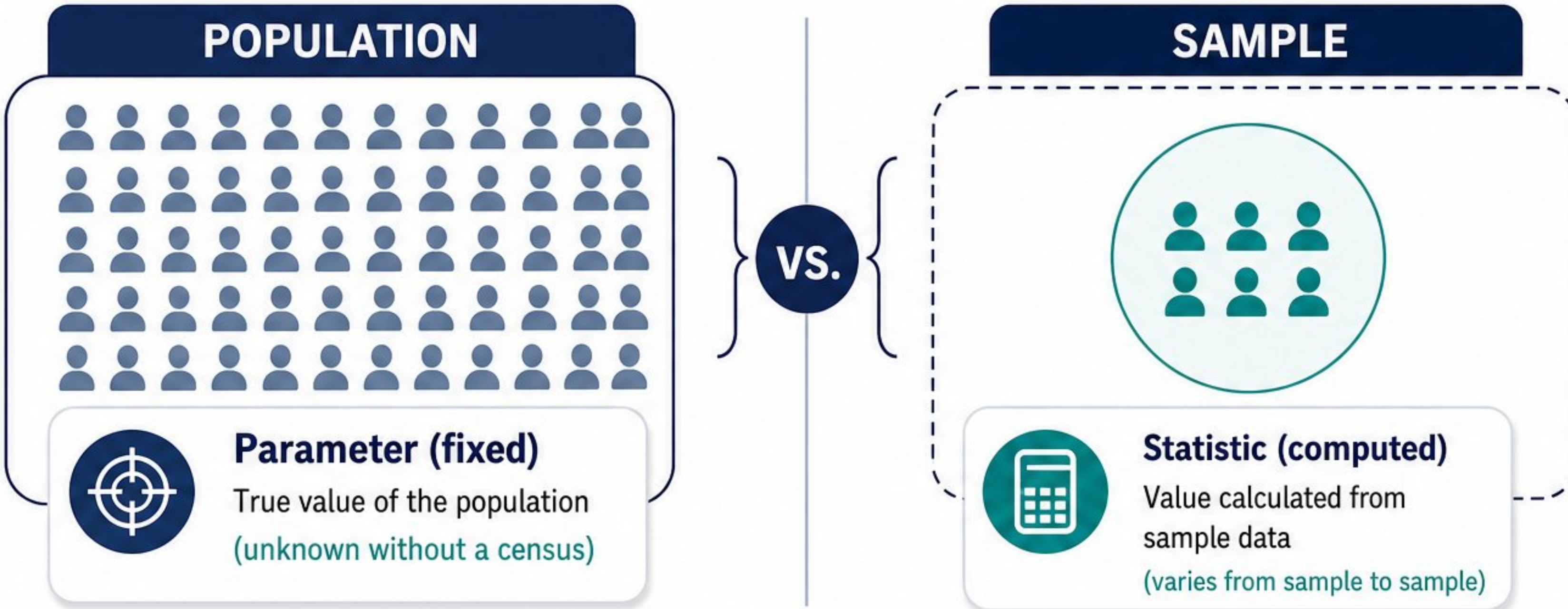
Read the notation and explain why the distinction drives inference

Roadmap: definitions, sampling variability, estimation, worked examples, pitfalls

By the end: name the parameter, choose an estimator, report uncertainty



What a Parameter Is, and What a Statistic Is



- **Parameter:** fixed numerical property of a population (mean, proportion, correlation)
- **Statistic:** number computed from sample data, such as $\overline{x}$ or $\hat{p}$
- Parameters are usually unknown without a full census
- Examples: true mean wait time, true voter support proportion, true mean wingspan
- Language habit: statistics in past tense, parameters in present tense

Notation: Greek Letters, Roman Letters, and Hats

Population (Parameters) Greek Letters	Sample (Statistics) Roman Letters
μ ↔	$\bar{x}$
σ ↔	s
σ^2 ↔	s^2
ρ ↔	r
β ↔	b

Ω
- Greek letters mark population parameters: $\mu, \sigma, \sigma^2, \rho, \beta$

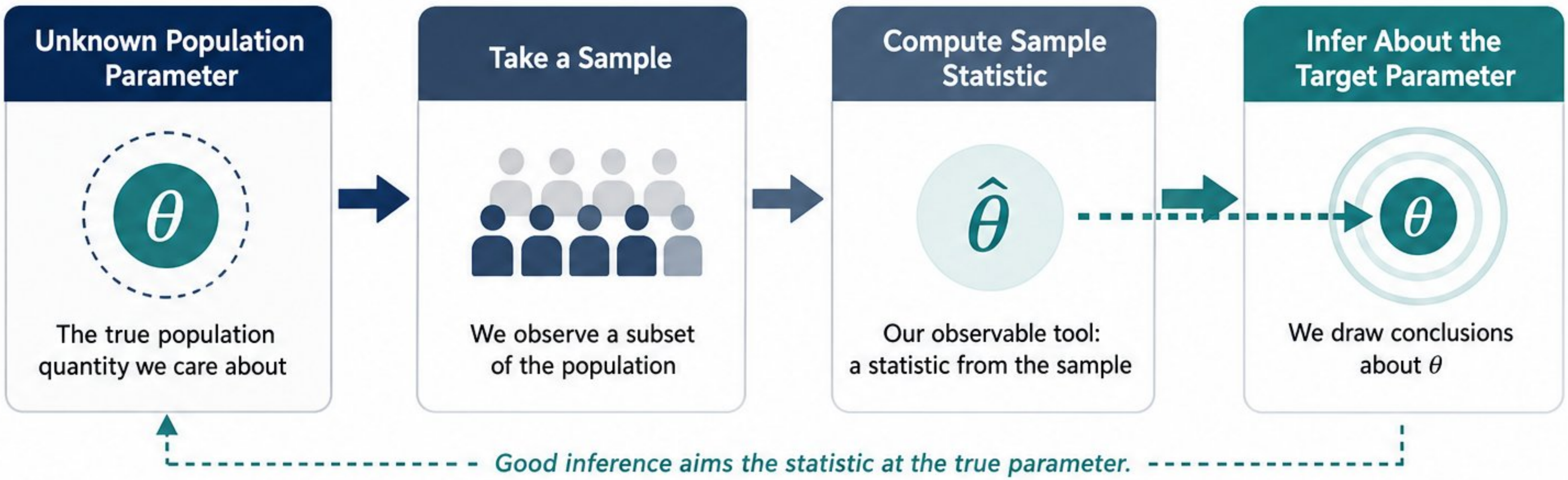
X̄
- Roman letters mark sample statistics: $\bar{x}, s, s^2, r, b$

- A hat means an estimate: $\hat{p}$ estimates p; $\hat{\beta}$ estimates β

- Estimator = the rule or formula; estimate = one sample's number

- Pitfall: lowercase p can mean proportion or p-value — always define it

Why Parameters Matter: The Targets of Inference



- Parameters describe the true population state; statistics are our observable tools
- Censuses are costly, slow, or impossible — samples are the practical route
- Inference uses sample statistics to answer population parameter questions
- Know which quantity you estimate; sample averages presented as truth mislead
- Descriptive questions about the sample need no parameter claim at all

A Catalogue of Common Parameters

- **Single-population:** mean μ, proportion p, variance σ^2, SD σ, median M
- **Relationship:** correlation ρ and regression coefficients β
- **Comparison:** difference in means $\mu_1 - \mu_2$ or proportions $p_1 - p_2$
- **Constraints by meaning:** proportions in $[0,1]$, variance ≥ 0, ρ in $[-1,1]$
- **Separate the parameter of interest** from nuisance parameters (e.g., error variance)

Parameter	μ Mean	p Proportion	σ^2 Variance	σ SD	M Median	ρ Correlation	β Regression coefficient
Meaning	Population average	Population proportion	Population variance	Population standard deviation	Middle value of ordered data	Linear association between two variables	Effect size in a regression model
Typical Range	$(-\infty, \infty)$	$[0,1]$	≥ 0	≥ 0	$(-\infty, \infty)$	$[-1,1]$	$(-\infty, \infty)$

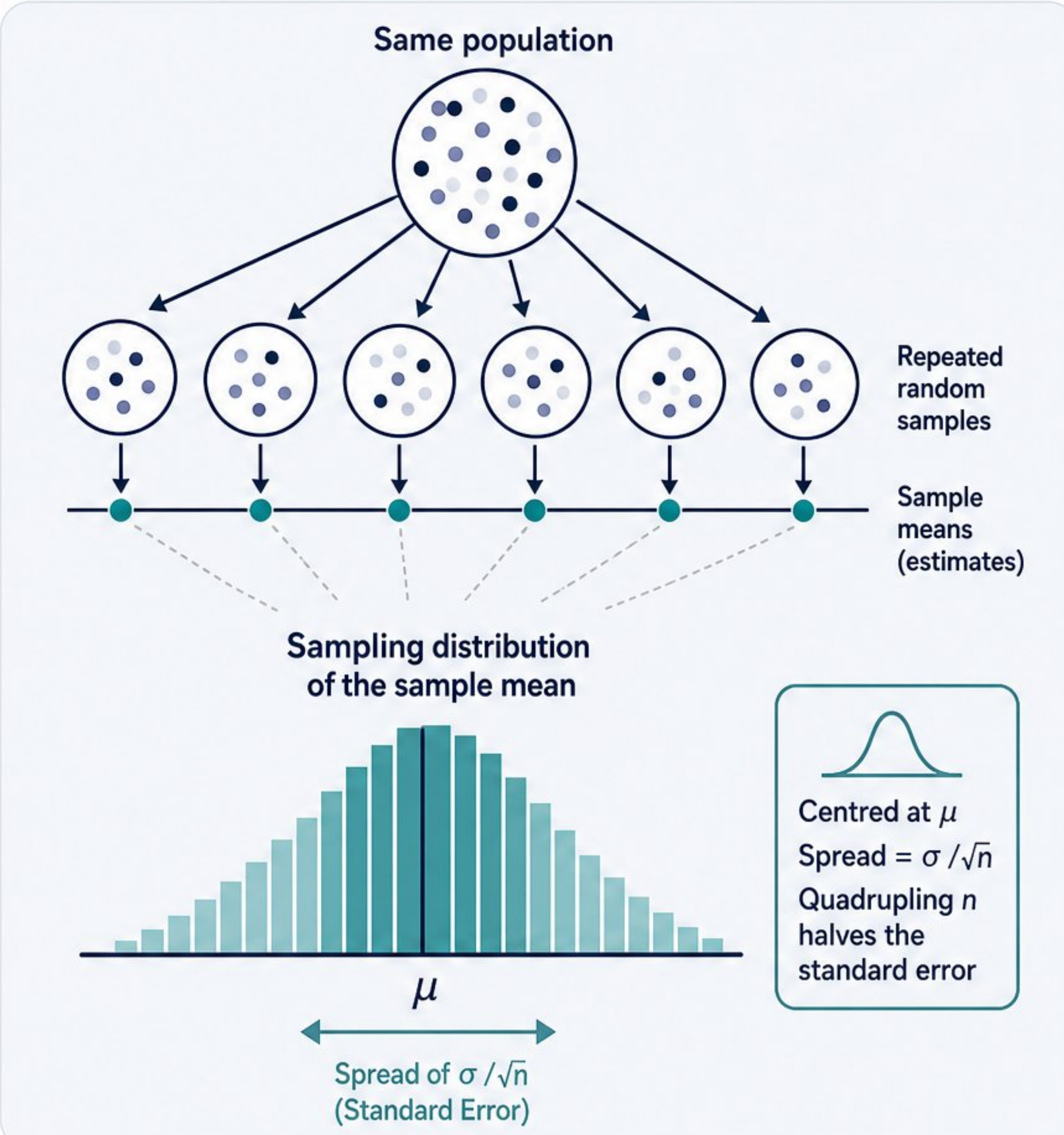
Notation Key

α **Greek symbols** (e.g., μ, σ, ρ, β)
Typically used for population parameters (unknown, fixed)

M **Roman symbols** (e.g., M, n)
Typically used for sample statistics or non-parameter quantities

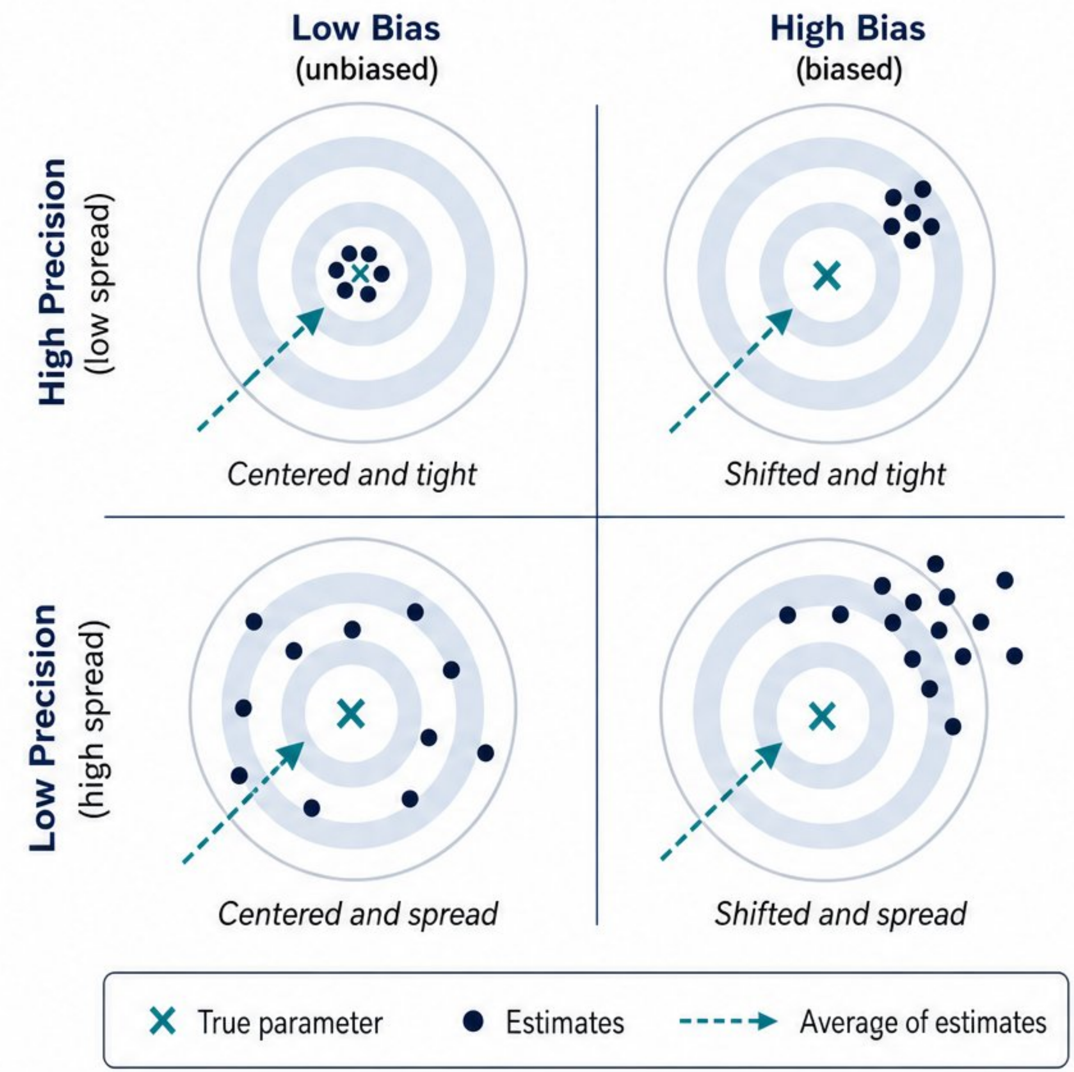
Sampling Variability: Why Estimates Differ

- Same population, sound sampling — yet every sample gives a different estimate
- Sampling error is expected variation, not a mistake
- **Sampling distribution:** values of a statistic across repeated samples
- **Standard error:** typical distance between an estimate and the parameter
- **Sample mean** centred at μ, spread $\sigma/\sqrt{n}$; quadrupling n halves the standard error

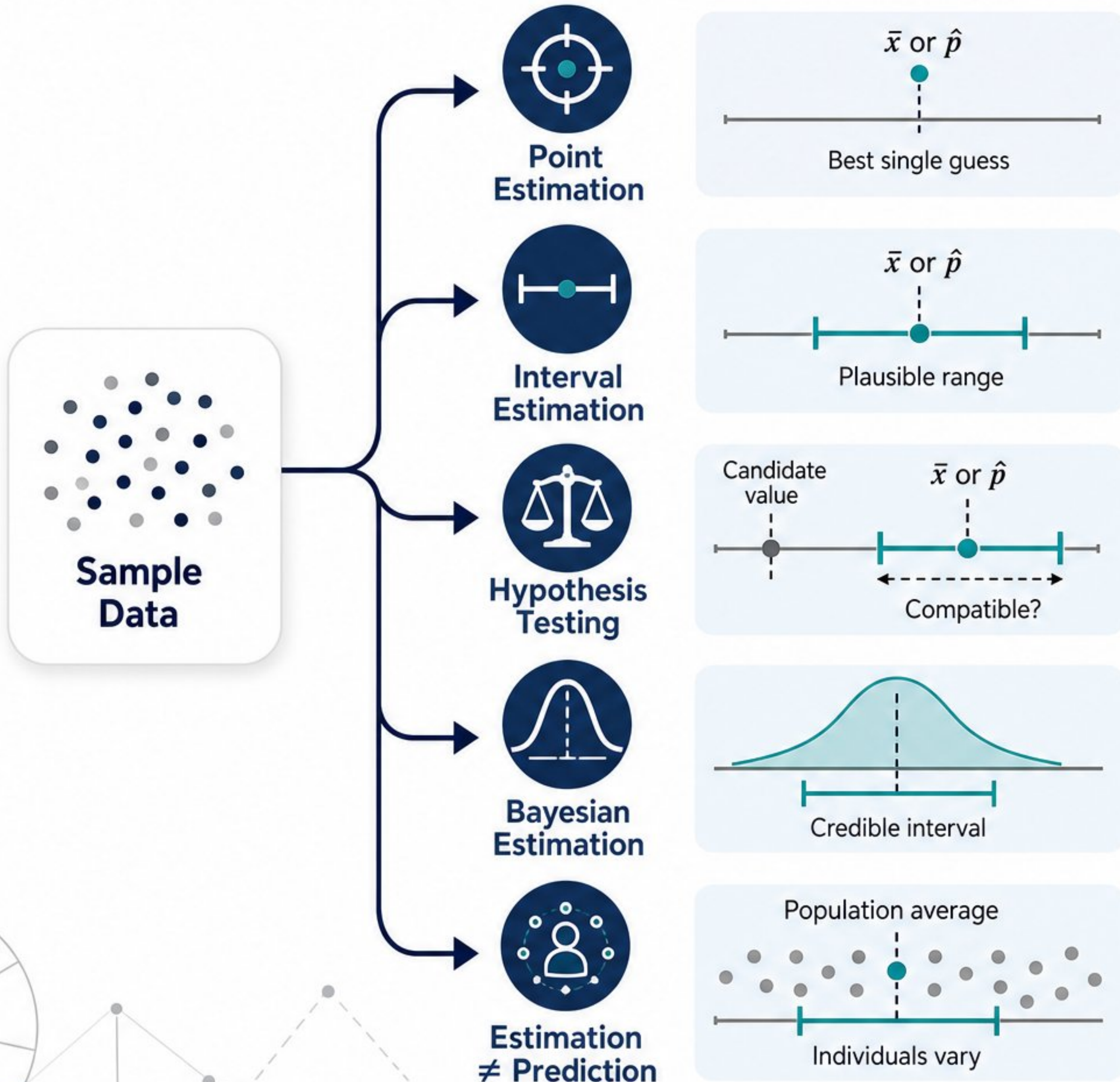


Bias, Precision, and the Large-Sample Picture

- **Bias:** average gap between estimator and true parameter
- **Precision:** how tightly estimates cluster around their own average
- **Consistency:** estimates converge to the parameter as *n* grows
- **Efficiency:** smallest variance among unbiased competitors
- **Bigger samples** fix randomness, not systematic bias



Estimating Parameters from Samples

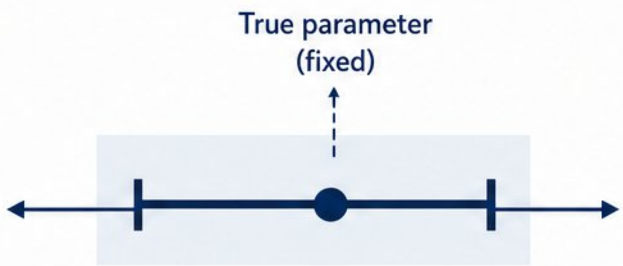


- **Point estimation:** one statistic, $\bar{x}$ or $\hat{p}$, as the best single guess
- **Interval estimation:** a range of plausible parameter values, width shows information
- **Hypothesis testing:** is a candidate parameter value compatible with the data?
- **Bayesian estimation:** posterior distribution and credible intervals express uncertainty
- **Estimation ≠ prediction:** a population average is not a future individual outcome

Interpreting Confidence and Uncertainty Correctly

MYTH

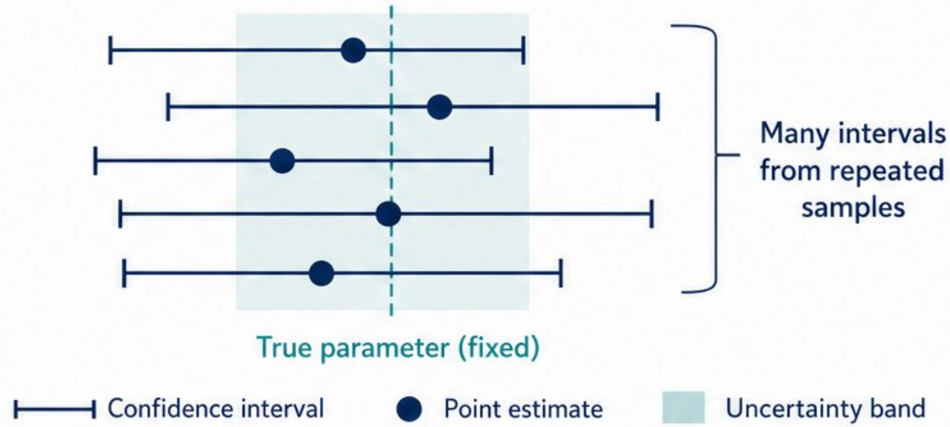
"95% probability the true parameter is in this interval"



Single interval

FACT

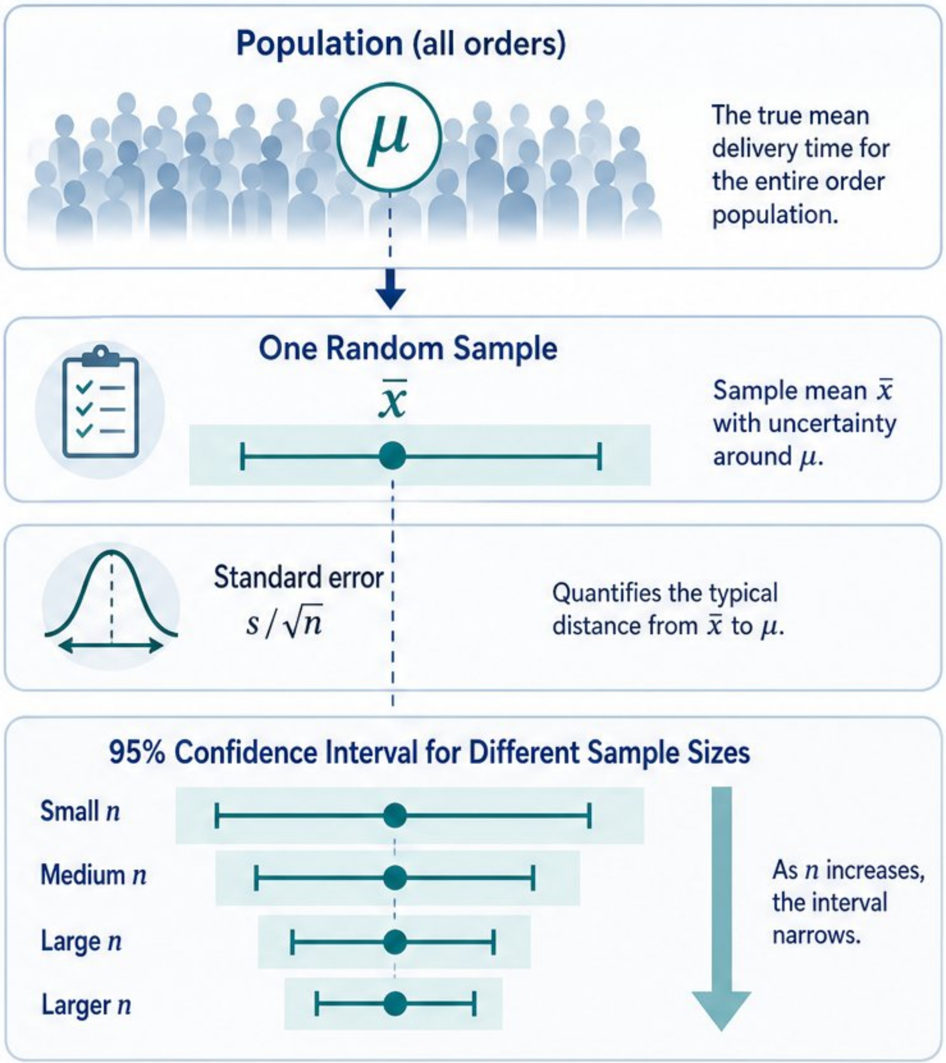
95% of intervals from the procedure will contain the true parameter



- 95% confidence describes the procedure, not the single interval
- Wrong: "95% probability the true parameter is in this interval"
- A 95% interval is not the range for 95% of future estimates
- Interval width measures precision; crossing zero is not the test
- Bayesian credible intervals allow direct probability statements, given priors

Worked Example: A Population Mean

- **Scenario:** average delivery time for all orders
- **Parameter μ:** true mean wait for the whole order population
- **Statistic $\bar{x}$:** sample mean from randomly sampled delivered orders
- **Standard error $s/\sqrt{n}$:** quantifies typical distance from μ
- **95% CI:** $\bar{x} \pm 1.96 \cdot s/\sqrt{n}$, with n and design assumptions stated
- **Larger n** narrows the interval: quadruple n to halve the standard error
- **Decision-maker reading:** "average wait is X minutes, plausibly A to B"

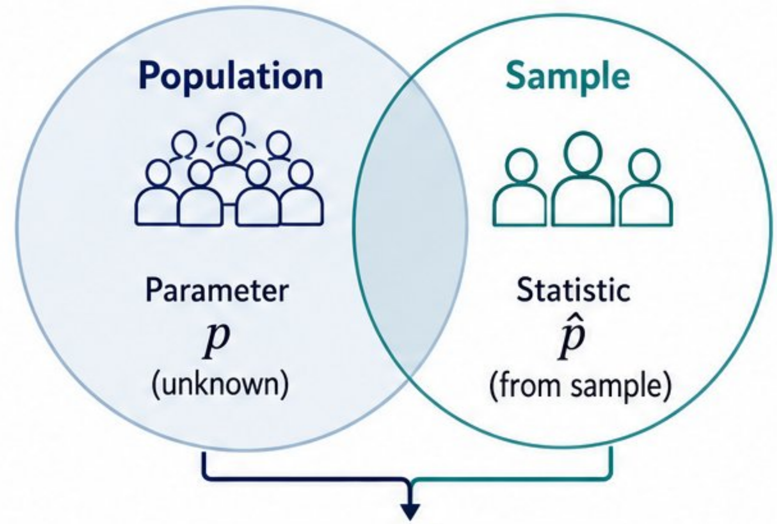


Takeaway: We estimate the population mean μ with a sample mean $\bar{x}$, reporting a 95% CI to express uncertainty.

Worked Example: A Population Proportion

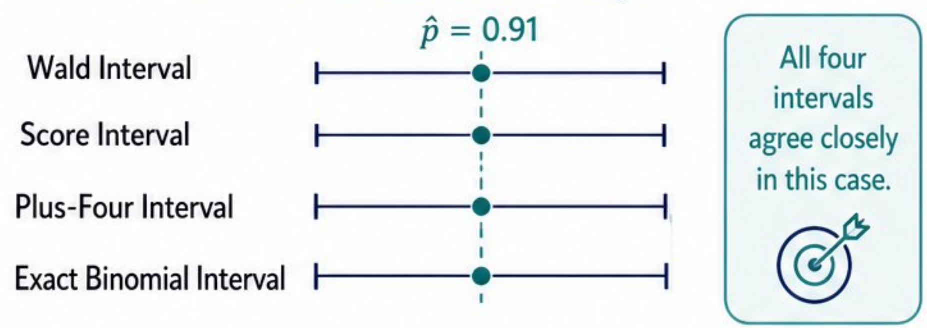
- **Scenario:** estimate support among all voters from a random sample.
- **Sample:** 1,987 respondents, 91% reporting support.
- **Name** the parameter p, the statistic $\hat{p}$, the estimate, and the standard error.
- Wald, score, plus-four, and exact binomial intervals agree closely here.
- **Check** validity conditions; state clearly which population the sample represents.
- **Flag** measurement bias: self-reported behavior can overstate true support.
- **Contrast parameter:** difference in means between two groups, $\mu_1 - \mu_2$.

Population vs. Sample



Estimate of p: $\hat{p} = 0.91$
from $n = 1{,}987$ respondents

Interval Methods Converge



Worked Example: Variance, Correlation, and Regression Parameters

- Population variance σ^2 and standard deviation σ target process variability
- Sample variance s^2 is unbiased for σ^2; s has slight downward bias
- Monte Carlo: simulate 1,000 samples, compare estimator spread to true values
- Correlation: population ρ versus sample r; regression slope β versus b
- Logistic regression coefficients describe population-level relationships with standard errors
- Sensitivity checks quantify how estimates shift with data or model changes

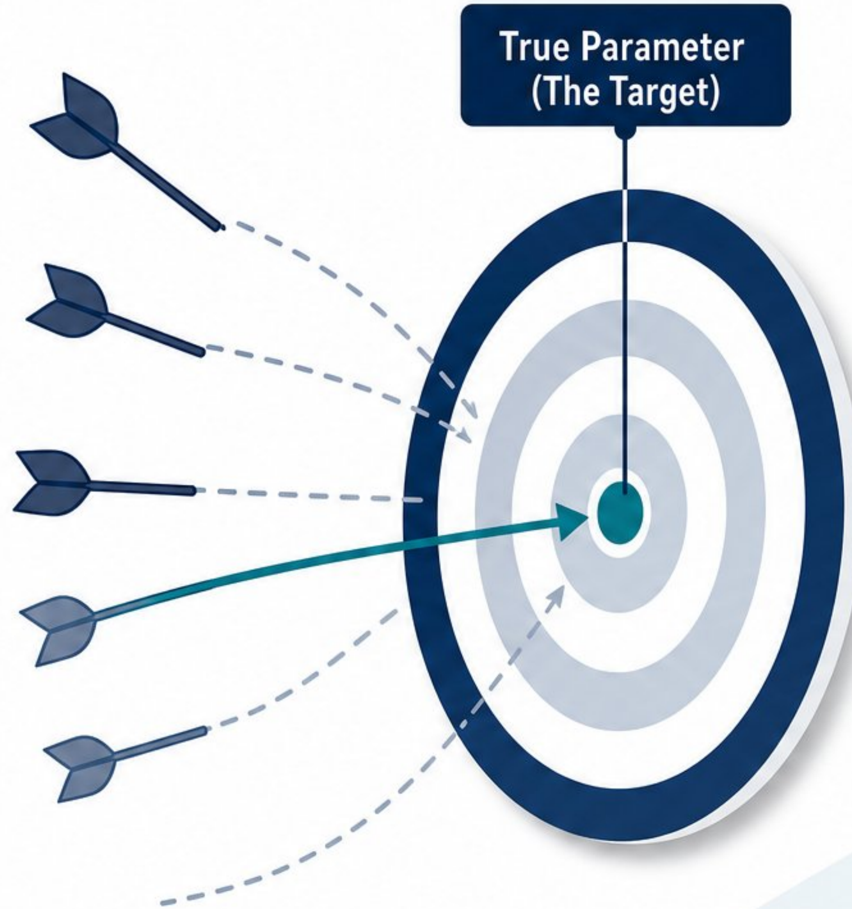
Monte Carlo: 1,000 Simulated Samples

Parameter	True Population (Value)	Sampling Distribution of Sample Estimator (1,000 Samples)	Estimator (Mean)
σ^2	Population variance σ^2	Spread of sample s^2	Average sample s^2
σ	Population standard deviation σ	Spread of sample s	Average sample s
ρ	Population correlation ρ	Spread of sample r	Average sample r
β	Population regression slope β	Spread of sample b	Average sample b
β (logit)	Logistic regression coefficients β (logit)	Spread of sample b (logit)	Average sample b (logit)

Center (true value) Spread (sampling variability) Sampling distribution of estimator

Common Misinterpretations and Pitfalls

- Treating a sample statistic as the fixed population parameter
- Reading a confidence interval as direct probability about the parameter
- Confusing statistical significance with practical importance
- Ignoring selection bias, nonresponse, and measurement error
- Overgeneralizing from convenience samples; misreading $1 - p$ as replication probability



Practical Guidelines for Reporting Parameters and Uncertainty

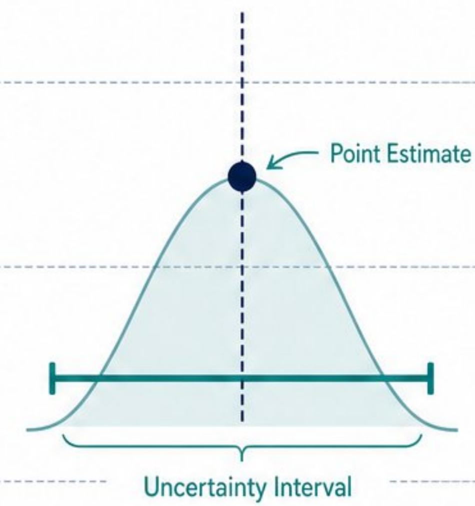
- Define target parameter and population before analysing
- Align sampling design and estimator with that parameter
- Report point estimate with standard error or interval
- State sample size, assumptions, and data sources clearly
- Communicate uncertainty numerically, visually, and in plain language
- Document code and limitations for reproducibility

REPORTING TEMPLATE

Parameter
Define the target parameter and the population of interest.

Point Estimate
Report the best estimate of the parameter.

Interval / Uncertainty
Provide a standard error or interval to quantify uncertainty.



Sample & Assumptions
State sample size, key assumptions, and data sources clearly.

Plain-Language Summary
Communicate what the estimate means and the uncertainty in simple terms.

> “We estimate the parameter to be around the point estimate. The true value is likely within the uncertainty interval, though it may be outside.”

Documentation
Document code and limitations to support reproducibility and transparency.

Key Takeaways and Knowledge Check

- Parameter: fixed population value; statistic: computed from a sample
- Sampling variability is unavoidable, so always report uncertainty
- Clear notation and terminology prevent confusion in analysis and teaching
- Judge estimators on bias, precision, and study design fit
- Knowledge check: classify parameters, statistics, estimates; fix confidence-interval misreadings

KNOWLEDGE CHECK

Practice and Apply

A Classify each statement as a PARAMETER, STATISTIC, or ESTIMATE.

1	The true average height of all adults in a country.	
2	The mean height calculated from a random sample of adults.	
3	Our best guess of the population mean based on the sample.	
4	The proportion of voters who will support a candidate.	
5	The standard deviation computed from the sample.	

B **Fix the confidence-interval misreading.**

Incorrect statement: “If a 95% confidence interval for μ is (L, U), then there is a 95% chance that μ lies between L and U.”

→ Rewrite the statement correctly.

NOTATION KEY

μ **Greek Letters** (Parameters) Describe fixed population values (constants).

$\bar{x}$ **Roman Letters** (Statistics / Estimates) Computed from samples or used as estimates of parameters.