BP201T · B.PHARM THEORY

BP201T Applied Biostatistics and Data Analytics for Pharmaceutical Sciences

Unit III — Sampling & Statistical Inference

Pills & Python Notes

From-zero teaching standard: every technical term is defined before use.

Syllabus coverage: Unit III (6 hours).

PILLS & PYTHON

Pharmacy • Data • AI • Healthcare

Instagram @pillsandpython · pillsandpython.com · YouTube @PillsAndPython

Chapter 3 — Sampling & Statistical Inference (Unit III, 6 hours)

3.1 Learning objectives

After studying this chapter, you will be able to:

1. Distinguish a population from a sample.
2. Name and compare sampling techniques used in clinical research.
3. Explain sampling error and bias with examples.
4. State the Central Limit Theorem in plain words and explain why it matters.
5. Construct and interpret a confidence interval.
6. Set up null and alternative hypotheses and walk through the hypothesis-testing framework.
7. Explain Type I and Type II errors, p-value, and statistical significance.
8. Demonstrate these calculations in Python with pharmaceutical data.

3.2 Population vs sample

Population = every member of the group you want to learn about (e.g., all diabetic patients in India). **Sample** = the smaller group you actually study (e.g., 200 patients in one trial). **Parameter** = a number describing the population (usually unknown, written with Greek letters like μ). **Statistic** = a number describing the sample (known, written with Latin letters like $\bar{x}$) — used to *estimate* the parameter.

Simple explanation: You cannot taste the whole pot of soup, so you taste one spoon — the pot is the population, the spoon is the sample, and you infer the pot's saltiness from the spoon. Inference = learning about the many from the few.

> **Analogy:** An election *exit poll* interviews a few thousand voters (sample) to predict crores of votes (population).

> **Key takeaway:** We study samples because populations are too big, costly, or impossible to measure fully — then we infer carefully.

3.3 Sampling techniques in clinical research

Technique	How it works	Strength	Weakness
Simple random	Lottery-style; everyone equally likely	Fair, no favouritism	Needs a full list of the population
Stratified	Divide into groups (strata, e.g., age bands), sample within each	Guarantees all subgroups represented	Needs to know the strata in advance
Systematic	Pick every kth person from a list	Easy to do	Risky if the list has a hidden pattern
Cluster	Randomly pick whole groups (e.g., 5 hospitals) and study everyone in them	Practical for spread-out populations	Less precise than random sampling
Convenience	Take whoever is available	Cheap and fast	Often biased — avoid for serious research

Pharmacy example: Testing a new antacid — stratified sampling by age ensures elderly patients (who use antacids most) are properly represented, not drowned out by young volunteers.

> **Key takeaway:** Good sampling is deliberate. Convenience sampling is the junk food of research — tempting, but unhealthy.

3.4 Sampling error and bias

Sampling error = the natural, unavoidable difference between a sample statistic and the true population parameter — it shrinks as the sample grows, and statistics can measure it. **Bias** = a *systematic* tilt that pushes results consistently in one direction — it does NOT shrink with bigger samples and cannot be fixed by statistics.

Worked illustration: A survey on drug side effects recruits only from a hospital's cardiology ward (convenience sample). Heart-drug side effects will be over-represented — that is **bias**. Even with 10,000 cardiology patients, the bias stays. If instead 200 patients are randomly picked from all wards, their average will differ slightly from the truth by pure chance — that is **sampling error**, and a bigger sample reduces it.

⚠️ **Terminology alert:** "Error" here does not mean a mistake — sampling error is natural wobble; bias is a tilted scale.

> **Key takeaway:** Bigger samples fix sampling error; only better *design* fixes bias.

3.5 The Central Limit Theorem (CLT) — the most important idea in statistics

Concept (plain words): Take many random samples from *any* population — even a wildly non-normal one — and the **sample means** will pile up into an approximately **normal** distribution, centred on the true population mean. Bigger samples → tighter, more normal pile.

Why it matters: It is the licence that lets us use normal-based tools (confidence intervals, t-tests) on real-world skewed data like hospital costs — provided samples are reasonably large (a common rule of thumb: n ≥ 30).

Standard error (SE): The SD of those sample means: **SE = σ ÷ √n**. It measures how much sample means wobble. Quadrupling the sample halves the wobble.

> **Worked example (illustrative):** Tablet weights have population SD σ = 5 mg. For samples of n = 25, SE = 5 ÷ √25 = 5 ÷ 5 = **1 mg**. For n = 100, SE = 5 ÷ 10 = **0.5 mg** — the mean of 100 tablets is a steadier estimate than the mean of 25.

> **Analogy:** One spoon of soup may mislead; the *average of 30 spoons* from different spots is trustworthy — CLT is the maths behind "average of averages settles down."

> **Key takeaway:** Sample means behave normally even when the raw data doesn't — that is why inference works.

3.6 Confidence intervals (CI) — a range of plausible truths

Definition: A 95% confidence interval is a range computed from the sample that, under repeated sampling, would capture the true population parameter 95% of the time. **Formula (large samples):** CI = $\bar{x} \pm 1.96 \times SE$, where $SE = s \div \sqrt{n}$.

> **Worked example (illustrative):** 36 tablets sampled: mean weight $\bar{x}$ = 500 mg, sample SD s = 3 mg.

- Step 1: $SE = 3 \div \sqrt{36} = 3 \div 6 =$ **0.5 mg**.
- Step 2: margin = 1.96 × 0.5 = **0.98 mg**.
- Step 3: 95% CI = 500 ± 0.98 = **(499.02, 500.98) mg**.
- Interpretation: we are 95% confident the true batch mean lies between 499.02 and 500.98 mg. Narrow interval = precise estimate.

⚠️ **Terminology alert:** "95% confident" does NOT mean "there is a 95% probability the true mean is inside *this* interval" — the true mean is fixed; it is the *method* that works 95% of the time.

Key takeaway: A CI reports both the estimate *and* its precision — wider means less certain.

3.7 Hypothesis testing framework

Concept: A formal "court trial" for a claim. **Null hypothesis (H_0)** = the boring default ("no effect / no difference"). **Alternative hypothesis (H_1 or H_a)** = the claim you are testing ("there IS an effect"). You collect evidence and decide whether to **reject H_0** — you never "prove" H_1, you only find H_0 untenable.

The five steps:

1. State H_0 and H_1.
2. Choose the significance level **α** (usually 0.05).
3. Collect data and compute the **test statistic** (e.g., t).
4. Find the **p-value**.
5. Decide: $p < \alpha \rightarrow$ reject H_0 ("statistically significant"); $p \geq \alpha \rightarrow$ do not reject H_0.

> **Worked example (illustrative):** A batch claims mean weight 500 mg. Sample: n = 25, $\bar{x}$ = 502 mg, s = 4 mg.

- Step 1: H_0: μ = 500 mg; H_1: μ ≠ 500 mg.
- Step 2: α = 0.05.
- Step 3: $t = (\bar{x} - \mu_0) \div (s \div \sqrt{n}) = (502 - 500) \div (4 \div 5) = 2 \div 0.8 =$ **2.5**. (df = 24.)
- Step 4: the critical t-value at α = 0.05 (two-sided, df 24) ≈ 2.064; our 2.5 exceeds it, so $p < 0.05$ ($p \approx 0.02$).
- Step 5: **Reject H_0** — the batch mean differs significantly from 500 mg; QC should investigate.

Key takeaway: Hypothesis testing = innocent (H_0) until proven guilty (by data, beyond α doubt).

3.8 Type I and Type II errors

	H_0 is actually TRUE	H_0 is actually FALSE
Reject H_0	**Type I error** (false alarm) — probability = α	Correct decision
Do not reject H_0	Correct decision	**Type II error** (missed signal) — probability = β

Definitions: Type I = seeing an effect that isn't there (convicting the innocent). Type II = missing a real effect (freeing the guilty). **Pharmacy examples:** Type I — approving a useless drug (patients pay, no benefit). Type II — rejecting a useful drug (patients lose a cure). Regulators fear Type I (hence α = 0.05); patients fear Type II.

Mnemonic: "**1** = false **1**-arm, **2** = missed **2**-ce" — or simply: *Type I cries wolf; Type II sleeps through the wolf.*

> **Key takeaway:** α controls false alarms; shrinking α raises the risk of misses (β) — a permanent trade-off.

3.9 p-value and statistical significance

Definition: The p-value is the probability of seeing data *this extreme or more extreme* **if H_0 were true**. Small p = the data would be surprising under H_0 = evidence against H_0.

Simple explanation: p = 0.02 means "if nothing were really happening, results like ours would occur only 2% of the time by luck." Because 2% < 5% (our α), we call it **statistically significant** and reject H_0.

⚠️ **Terminology alert — what p is NOT:** p is not P(H_0 is true), not P(H_1 is true), and not the size or importance of the effect. A tiny p-value with a trivial effect (e.g., 0.1 mg weight difference in 10,000 tablets) is statistically significant but practically meaningless — always report the *effect size* alongside p.

> **Analogy:** p-value is the *court's doubt-meter*: p < 0.05 = "beyond reasonable doubt" → convict (reject H_0).

> **Key takeaway:** p < α → reject H_0. But statistical significance ≠ clinical importance.

3.10 Demonstration in Python (illustrative data)

Snippet 1 — confidence interval by hand formula:

```
import numpy as np

weights = np.array([502, 498, 505, 500, 503, 497, 501, 499, 504, 500,
                    502, 499, 501, 503, 500, 498, 504, 501, 500, 502])
xbar = np.mean(weights)         # sample mean
s = np.std(weights, ddof=1)     # sample SD (divide by n-1)
se = s / np.sqrt(len(weights))  # standard error
margin = 1.96 * se              # 95% margin
print("95% CI:", (xbar - margin, xbar + margin))
```

Line-by-line: compute the mean, the sample SD (`ddof=1`), divide by √n for the SE, multiply by 1.96 for the 95% margin, and print the interval. Sample output (computed by running the code):

```
95% CI: (499.99, 501.91)
```

Interpretation: We are 95% confident the true batch mean lies between 499.99 and 501.91 mg — a tight interval around the 500 mg target, so the batch looks fine.

Snippet 2 — one-sample t-test with SciPy (standard pattern):

```
from scipy import stats

t_stat, p_value = stats.ttest_1samp(weights, popmean=500)
print("t =", t_stat, " p =", p_value)
```

Line-by-line: `stats.ttest_1samp(weights, popmean=500)` tests H_0: true mean = 500 against the sample; it returns the t-statistic and the p-value. If the printed $p < 0.05$, reject H_0 — the batch mean differs significantly from 500 mg.

> **Key takeaway:** Python does the arithmetic; *you* supply the hypotheses and the interpretation.

3.11 Pharmacy relevance (Unit III)

- **Clinical trials:** Every "drug A beats placebo ($p < 0.05$)" headline is hypothesis testing.
- **QC:** Batch release uses confidence intervals and t-tests against specifications.
- **Research papers:** You cannot read a single clinical paper without CIs and p-values.

3.12 Common confusions

1. **Statistic vs parameter:** statistic describes the sample ($\bar{x}$, s); parameter describes the population (μ, σ).
2. **Sampling error vs bias:** error shrinks with n; bias doesn't.
3. **"Accept H_0"** — wrong phrase. We *fail to reject* H_0; absence of evidence is not evidence of absence.
4. **p-value** is not the probability that H_0 is true.

3.13 Common mistakes

- Using convenience samples and generalising boldly.
- Reading "95% CI" as "95% probability the mean is inside."
- Confusing statistical significance with practical importance.
- Running a t-test on heavily skewed tiny samples where CLT can't help.

3.14 One-minute revision (Unit III)

- **Population** (all) vs **sample** (studied few); **parameter** (μ, unknown) vs **statistic** ($\bar{x}$, known).
- Sampling: simple random, **stratified** (subgroups), systematic, cluster, convenience (biased — avoid).
- **Sampling error** shrinks with n; **bias** is systematic and doesn't.
- **CLT:** sample means ≈ normal, centred on μ; **SE = σ ÷ √n**.
- **95% CI = $\bar{x}$ ± 1.96 × SE** — estimate + precision.
- H_0 (no effect) vs H_1; $p < \alpha$ (0.05) → **reject H_0**.
- **Type I** = false alarm (α); **Type II** = missed signal (β).
- **p-value** = P(data this extreme | H_0 true); small p = evidence against H_0.

3.15 Practice questions (Unit III)

MCQs

1. Which sampling method guarantees representation of age subgroups?

a) Convenience b) Stratified c) Systematic d) Cluster

2. Increasing sample size reduces:

a) Bias b) Sampling error c) Both d) Neither

3. A Type I error is:

a) Missing a real effect b) False alarm — rejecting a true H_0 c) Using the wrong test d) A calculation mistake

4. p = 0.03 with α = 0.05 means:

a) Accept H_0 b) Reject H_0 c) The effect is large d) H_0 is 3% likely true

Short questions

5. Distinguish sampling error from bias with one example each.
6. State the Central Limit Theorem and explain why it justifies using t-tests on skewed data.
7. Interpret: "95% CI for mean weight loss: (2.1, 3.4) kg." What does it tell a clinician?

Long / application-based

8. (Illustrative) A syrup claims 100 mg/5 mL. Sample n = 36: $\bar{x}$ = 99.2 mg, s = 3 mg. Test H_0: μ = 100 at α = 0.05 showing all five steps, and write the QC conclusion.
9. Explain with a 2×2 table when a Type I vs Type II error is more dangerous in drug approval, and why regulators fix α at 0.05.
10. Write Python code to compute the 95% CI for a list of dissolution values; explain each line and interpret the output for batch release.

Unit III Infographic — One-Minute Concept Map

Population (all) vs **sample** (studied few); **parameter** (μ, unknown) vs **statistic** ($\bar{x}$, known).

Sampling: simple random, **stratified** (subgroups), systematic, cluster, convenience (biased — avoid).

Sampling error shrinks with n; **bias** is systematic and doesn't.

CLT: sample means ≈ normal, centred on μ; **SE = σ ÷ √n**.

95% CI = $\bar{x}$ ± 1.96 × SE — estimate + precision.

H_0 (no effect) vs H_1; $p < \alpha$ (0.05) → **reject H_0**.

Type I = false alarm (α); **Type II** = missed signal (β).

p-value = P(data this extreme | H_0 true); small p = evidence against H_0.

Combined References

1. Lane, D.M., *Introduction to Statistics*. Houston: Rice University. Available at: https://onlinestatbook.com/2/index.html
2. Walters, S.J., Campbell, M.J. and Machin, D., *Medical Statistics: A Textbook for the Health Sciences*. Hoboken, NJ: Wiley-Blackwell.
3. Rowe, P., *Essential Statistics for the Pharmaceutical Sciences*. London: Pharmaceutical Press.
4. Bolton, S. and Bon, C., *Pharmaceutical Statistics: Practical and Clinical Applications*. New York: Informa Healthcare.
5. De Muth, J.E., *Basic Statistics and Pharmaceutical Statistical Applications*. Boca Raton, FL: CRC Press.