



Biostatistics 5 — Case Study Report

Effect of Drug Dose \times Time on Protein Expression

SupBiotech — TD/TP Session

Name(s): _____

Submission Instructions

- Responses must follow sections **1–7** below. Do **not** copy the prompts; write concise, numbered answers only.
- Use a research style (methodology \rightarrow results \rightarrow interpretation); be specific and data-driven.
- All figures appear **at the end** (Section 9), each with a short caption.
- **Full R code** used to generate results and figures must be included in Section 10 (Code Appendix). Replace all ?? placeholders with your own variables, tests, and paths.
- The **regression model is obligatory**. A comparison between test-based approaches and the model is required.
- Indicate group contributions if applicable (who did what).

1. Data Exploration

- Dataset size: number of observations and variables.
- Missingness: how many missing values and where; handling choice and justification.
- Structure summary (e.g., counts across design cells).
- Descriptives: mean, standard deviation, n (per relevant grouping).
- One descriptive figure (box/violin/histogram) referenced in text.

Response 1: _____

2. Assumption Checks

- Distributional check(s) and result(s).
- Variance check(s) and result(s).
- Decision: are assumptions sufficiently met? If not perfect, does the CLT (given sample sizes) justify parametric methods?
- Brief justification referencing your results (not generic statements).

Response 2: _____

3. Test or Model Selection

- At least two *different* comparison approaches for hypotheses (e.g., two-level and multi-level), and the **regression model specification** (obligatory).
- Rationale for each choice (why this approach is appropriate for the question and data).
- Alternatives considered and why they were not selected.
- Any transformation attempted and the reason.

Response 3: _____

4. Results

- Results from each chosen test: effect(s), p -values, confidence intervals (where relevant).
- Model results: estimated effects/coefficients, significance, and any interaction.
- **Comparison of methods:** succinctly contrast outcomes from the test-based approaches vs the regression model. Note convergence or divergence of conclusions and why this may occur.
- **Most significant/clinically relevant finding:** identify and justify.

Response 4: _____

5. Post-hoc Analysis

- If applicable, post-hoc procedure(s) and key pairwise differences.
- How post-hoc findings align with main results.

Response 5: _____

6. Biological Interpretation

- Interpretation of supplement effects, time effects, and interaction in biological terms.
- Brief discussion of plausible mechanisms/implications for enzyme activity.

Response 6: _____

7. Reflection

- Most difficult aspect of the analysis and how it was addressed.
- What worked well; limitations of the current approach.
- What would change with more time/data (e.g., additional diagnostics, alternative models).
- Key insight gained from comparing different approaches.

Response 7: _____

8. Comparison Matrix & Visual Summary

- Provide a tagged matrix of comparable statistics across approaches (e.g., p -values for two-level, multi-level, and model).
- Include a short (2–5 lines) interpretation of similarities/differences among approaches and the model.

Response 8: _____

9. Figures (all plots together)

Insert figures produced by the analysis. Each must have a short caption. Recommended minimum set:

- Two two-level comparison plots (e.g., box or violin for two distinct subsets).
- One multi-level comparison plot.
- One mean \pm 95% CI plot.
- One regression/model view (obligatory).
- One comparison visual (e.g., bar/heatmap) for the matrix in Section 8.