

BCB744 Biostatistics Exam Rubric (2025)

Smit, A. J.

2025-05-31

General Structure of the Rubric

Each Task is evaluated under the following axes:

1. Technical Accuracy (50%)
2. Depth of Analysis (20%)
3. Clarity and Communication (20%)
4. Critical Thinking (10%)

Each subcomponent is marked on a 0–100 scale, then scaled to its proportion of the task weight. For example, Task 5 is worth 30% of the total mark, so a sub-question like 5.1 (one of five) may contribute up to 6% if evenly weighted.

Task 1: Initial Processing [10%]

Weight within task:

- 1.1 Extraction and Restructuring (50%)
- 1.2 Conversion and Summarisation (50%)

Rubric:

- Technical Accuracy (50%)
 - Correct unpacking of NetCDF variables (names, dimensionality): 15%
 - Time conversion handled correctly (POSIX timestamps): 10%
 - Data reshaped into appropriate long format: 15%
 - Presence of appropriate columns (year, quarter, etc.): 10%
- Depth of Analysis (20%)
 - Efficient use of methods (e.g. `hyper_tibble()` or `expand.grid()` vs brute loops): 10%
 - Use of Cartesian indexing or equivalent vectorised operation: 10%
- Clarity and Communication (20%)
 - Code is readable, well-commented: 10%
 - Summary of the resulting data structure shown and interpretable: 10%
- Critical Thinking (10%)
 - Indicates understanding of spatial \times temporal structure and mentions NA implications: 10%

Task 2: Exploratory Data Analysis [10%]

2.1 Weighted Mean Time Series

- (1) Weighted mean across time: 15%
- (2) Time series for 100 pixels: 15%

2.2 Summary Statistics:

- (1) Descriptive stats: 20%
- (2) Visualisations: 20%
- (3) Interpretation: 20%

2.3 Observation Density Map: 10%

Rubric:

- Technical Accuracy (50%)
 - Proper handling of weights and NA filtering: 10%
 - Correct aggregation logic (quarter, pixel, etc.): 10%
 - Appropriateness of visualisation syntax and ggplot conventions: 10%
 - Use of statistical descriptors (mean, sd, skew, etc.) correctly: 10%
 - Map projection/geodesic coordinates and section overlay accuracy: 10%
- Depth of Analysis (20%)
 - Commentary on skewness, kurtosis, and statistical implications: 10%
 - Recognition of seasonal/temporal signals in plots and stats: 10%
- Clarity and Communication (20%)
 - Plot labels, axes, titles intelligible and precise: 10%
 - Logical narrative supporting visualisations/statistics: 10%
- Critical Thinking (10%)
 - Justification of metric choices, handling of anomalous years: 5%
 - Suggestions of ecological explanations (e.g., photoperiod, storminess): 5%

Task 3: Inferential Statistics I [20%]

Weight within task:

- (1) Hypotheses: 10%
- (2) Model selection and justification: 20%
- (3) Assumption testing: 20%
- (4) Result interpretation and diagnostics: 50%

Rubric:

- Technical Accuracy (50%)
 - Correct use of linear model and specification (additive, no interaction): 20%
 - Explicit assumptions tested (normality, homogeneity): 10%
 - Proper model diagnostics and visual checks: 10%
 - Use of correct significance thresholds and p-value interpretation: 10%
- Depth of Analysis (20%)
 - Justification for using aggregate means vs raw data: 10%
 - Consideration of alternative models (e.g., GAMs): 10%
- Clarity and Communication (20%)
 - Hypotheses stated cleanly, concisely: 10%
 - Figure/Table references integrated smoothly in the narrative: 10%
- Critical Thinking (10%)
 - Recognition of model limitations and implications (e.g. low R^2): 10%

Task 4: Spatial Assignment [10%]

4.1 Section Assignment: 5%

4.2 Bioregion Assignment: 5%

Rubric:

- Technical Accuracy (50%)
 - Correct application of Haversine formula or great-circle logic: 20%
 - Accurate section_id assignment: 10%
 - Bioregion mapping via join or merge: 10%
 - Correct data columns preserved/renamed: 10%
- Depth of Analysis (20%)
 - Efficiency of matching routine (e.g., mapply() or vectorised join): 10%
 - Consideration of spatial boundaries (e.g., limiting to section 1–22): 10%
- Clarity and Communication (20%)
 - Annotated code, explanation of proximity logic: 10%
 - Output (head(), summary(), tail()) shows assignment integrity: 10%
- Critical Thinking (10%)
 - Considers effect of section resolution or mapping error: 10%

Task 5: Inferential Statistics II [30%]

Each sub-task contributes approximately 6% unless reweighted explicitly.

Rubric per sub-task (5.1–5.5):

- Technical Accuracy (50%)
 - Model type (ANOVA, LM, ANCOVA) appropriate: 15%
 - Correct test execution (summary, diagnostics): 15%
 - Assumptions evaluated, violations addressed: 10%
 - Non-parametric alternative proposed when appropriate: 10%
- Depth of Analysis (20%)
 - Explicit rationale for model choice: 10%
 - Discussion of structure in data (nesting, lack of interaction): 10%
- Clarity and Communication (20%)
 - Hypotheses clearly and formally stated: 10%
 - Visualisations appropriately labelled and explained: 10%
- Critical Thinking (10%)
 - Insight into ecological implications of findings (e.g., BMP trend): 10%

Add 1–2 bonus marks if:

- Multicollinearity (e.g., VIF) or autocorrelation (e.g., DW test) is discussed
- Advanced diagnostics (e.g., Breusch–Pagan, TukeyHSD) are used correctly

Task 6: Write-up [10%]

Rubric:

- Technical Accuracy (50%)
 - Consistent reference to previous results, correct figure/table interpretation: 25%

- Accurate paraphrasing of statistical results: 15%
 - Adherence to 2-page length limit, integration of material: 10%
- Depth of Analysis (20%)
 - Rich synthesis across Tasks 2–5, not isolated repetition: 10%
 - Conceptual connection of seasonality, trend, and spatial heterogeneity: 10%
- Clarity and Communication (20%)
 - Coherent scientific writing style, flowing paragraph structure: 10%
 - Effective integration of figure references and literature: 10%
- Critical Thinking (10%)
 - Limitations clearly acknowledged and reflected on: 5%
 - Forward-looking ecological insight or recommendation offered: 5%