Properties of Alternative Dirichlet-Multinomial Forms

For Fitting Composition Data

Goal

- Evaluate the relative performance of three alternative composition likelihood models:
    - Multinomial – the large sample limit of the Saturated Dirichlet-multinomial
    - Linear Dirichlet-multinomial
    - Saturated Dirichlet-multinomial
- Support better selection of composition likelihoods in integrated fishery stock assessment models.
- Use function-level diagnostics to assess estimator quality over the probability simplex for observed composition data from a simulated fishery system.

Problem

- Traditional modeling of age or size composition data often fails to account for:
    - **Overdispersion** beyond the multinomial assumption.
    - **Residual correlation** between composition bins.
    - **Intracluster correlation** among composition data samples.
- This can lead to biased estimates, underestimated uncertainty, and poor model fit in stock assessments that use the multinomial distribution to model composition data.

Solution

- Implement a **functional analysis framework in C++ and R** to:
    - Sample observed proportions from a simulated fishery system over appropriate subsets of the **probability simplex**.
    - Evaluate each distribution using:
        - Log-likelihood values
        - Stochasticity and variance metrics
        - Sensitivity diagnostics from **gradients and Hessians**
        - Residual correlation structure
        - Effective sample size estimation
        - Bias in recovered proportions
        - Convergence and speed metrics
        - Accuracy and precision of predictions.
- Provide **practical guidance** on the selection of composition model likelihoods based on robustness, informativeness, and computational stability.