An evaluation of alternative binning approaches for composition data in integrated stock assessments

# Main questions

1. What is the optimal strategy for binning length composition data, given a life history.
2. What is a rule of thumb for setting the robustification constant?
3. How and when should tail compression be used?
4. Is model selection a reliable way to determine data bins?

# Main dimensions and cases

1. Data quantity/quality
   1. Sample sizes, timing, etc. of age, length, and indices.
2. Length bin definitions.
   1. Equally-spaced bins, of varying width.
   2. Bins spaced such that all bins have equal (on average) samples
   3. Unequal bins, based on biology or fishery property.
3. Robustification constant
   1. Ranges from 0 to some fraction of smallest observed proportion.
4. Tail compression
   1. Based on the observed data
   2. Arbitrary value, from a rule of thumb.

# Biological assumptions

1. Natural mortality fixed @ true value.
2. BH recruitment function, fixed @ true value?
3. Growth parameters constant in OM but estimated in EM.
4. Weight/Length/Maturity fixed @ true value?
5. R0 estimated

# Fishery assumptions

1. OM fishery selectivity time-varying? EM selectivity estimated as constant.
2. Catchability coefficients estimated for fishery and survey (set to 1 in OM)

# Metrics

Relative error (RE) and median absolute relative error (MARE) for: management targets (MSY, depletion, etc.), and growth parameters

# Potential tables & figures

1. Figure of robustification constant vs. MARE of metrics
2. Figure of constant bin width (or # bins) vs. MARE of metrics and growth parameters
3. Figure of tail compression vs. MARE of metrics