**Observation for length frequency data**

Need to generate an age-length key (ALK) from the growth of length by age, which also specifies the distribution of lengths around the mean for each age. The ALK is applied to the population-by-age to get the population by age and length. We sum over each length class to get the population-by-length class, and then divide this through by the total population to get an expected length frequency. The expected length frequency is then used in the log-likelihood with the sample length frequency.

**Main routines**:

*Cum\_normal*: calculates the probability of lengths being less than or equal to given length boundary using an approximation to the cumulative normal curve. (line 80) Subtraction of next door values gives the probability of fish in each length class. Special case when there is a plus-group.

*Get\_age\_length*: loops through all ages and generates the age-length key (calls *Cum\_normal*) (line 182)

Code in obs\_length\_code1.cpp tests the above functions against validation checks made in R using brute force random sampling. *Cum\_normal*: tested both lognormal and normal distributions (lines 264 – 280). *Get\_age\_length*: tested with 10 ages and given numbers-at-age (Line 281 onwards).

Text file, "obs\_length\_results.txt", contains comparison of C++ results with R validation results (Line 347 onwards).

**Length observations.**

1. Line 303: Lower boundaries for each length class are set up and log version setup, if needed.
2. Line 294: Get growth mean lengths and standard deviation for each age
3. Line 312: get age-length matrix (*age\_length* variable) from the growth and its distribution (before applying numbers-at-age). Age-length key (ALK)
4. Line 318: Get the population by age & length matrix (rows are ages, columns are length classes) by multiplying each row by its population for that age (*pop\_age\_length* variable).
5. Line 328: Expected length frequency, get numbers-at-length by adding each column in *pop\_age\_length*, then for the length frequency, divided by the population total in *pop\_age\_length*.
6. Expected length frequency is used in the log-likelihood function with the sample frequency