4.5.5 - Uniformity Measurements

The Flux Simulator provides some statistical indicators to measure the uniformity of the read distribution along a transcript produced by *in silico* sequencing.

Fraction Covered

The fraction of a transcript that is covered by reads reflects its expression and the degree of of coverage fluctuation provoked by biases.

Unknown macro: 'mathinline'   where   Unknown macro: 'mathinline'   is the length of the transcript,

and the sign function

Unknown macro: 'mathblock'

indicating whether a position is covered by at least one sequenced read.

Chi-square statistics (X^2)

Pearson's chi-square can be used to test the goodness of fit of a given sample to a theoretical distribution. Given a transcript of length

Unknown macro: 'mathinline'   and coverage   Unknown macro: 'mathinline'   at position

Unknown macro: 'mathinline' , the test statistic is defined as follows:

Unknown macro: 'mathblock'

where   Unknown macro: 'mathinline'   is the average coverage along the molecule.

Coefficient of variation (CV)

A CV is defined as the ratio of the standard deviation and the standard deviation of a probability distribution:

Unknown macro: 'mathinline' , with standard deviation   Unknown macro: 'mathinline'   and mean

Unknown macro: 'mathinline'

of transformed coverage values   Unknown macro: 'mathinline'

Latter Anscombe transformation of coverage values has been proposed [Hansen et al. 2010] under the assumption that the distribution of reads along a transcript follows a Poisson distribution, which is to be transformed to a Gaussian distribution.