### 1 Complete protein-spectrum-matches without PTMs.

(a) Complete protein n terminal residue frequencies using modified protein list. The sum of all frequencies is 1.

(b) The table probabilities in peak\_begin and peak\_end are summed up to obtain precursor probabilities.

(c) Conditional probabilities are computed based on the ratio (probability above threshold) / (total probability)

(d) Count number is the number of all modified proteoforms with a similar precursor mass.

(e) adjust factor is 1.0

(f) e-value is conditional\_probability \* count\_number \* adjust\_factor, p-value is 1 - (1-conditional\_prob)^(count\_number \* adjust\_factor)

\*\*The current implementation is correct.\*\*

### 2 Prefix protein-spectrum-matches without PTMs.

(d) Count number is the number of modified prefixes with a similar precursor mass

\*\*The current implementation is correct.\*\*

### 3 Suffix protein-spectrum-matches without PTMs

(a) Use peptide n terminal residue frequency list (not protein n terminal residue frequency list)

(d) Count number is the number of suffixes with a similar precursor mass.

\*\*The current implementation is correct.\*\*

### 4 Internal protein-spectrum-matches without PTMs

(a) Use peptide n terminal residue frequency list (not protein n terminal residue frequency list)

(d) Count number is the number of internals with a similar precursor mass.

\*\*The current implementation is correct.\*\*

### 5. Complete protein-spectrum-matches with one PTM.

(a) Complete protein n terminal residue frequencies using modified protein list. The sum of all frequencies is 1.

(b) The allowed shift for a PTM is [0, average\_residue\_length]. Follow the methods described in the spectral probability paper.

(b) The table probabilities in peak\_begin and peak\_end are summed up to obtain precursor probabilities.

\*\*(c) Conditional probabilities are computed based on the ratio (probability above threshold / shift probability \* peak\_width). The reason for multiplying peak\_width is the error tolerance.\*\*

\*\*(d) Conditional probabilities is normalized by \* K (0.55)\*\*

(e) Count number is the number of modified proteoforms.

(f) adjust factor is 1.0

(g) e-value is conditional\_probability \* count\_number \* adjust\_factor, p-value is 1 - (1-conditional\_prob)^(count\_number \* adjust\_factor)

### 6 Prefix protein-spectrum-matches with one PTM.

(e) Count number is the number of modified prefixes \* \*\*PREFIX\_SUFFIX\_ADJUST (0.693)\*\*

### 7 Suffix protein-spectrum-matches with one PTM.

(e) Count number is the number of modified suffixes \* \*\*PREFIX\_SUFFIX\_ADJUST (0.693)\*\*

### 8 Internal protein-spectrum-matches with one PTM.

(e) Count number is the number of modified internals \* \*\*INTERNAL\_ADJUST (0.508) --> (0.05)\*\*

### 9. Complete protein-spectrum-matches with two PTM.

(a) Complete protein n terminal residue frequencies using modified protein list. The sum of all frequencies is 1.

(b) The allowed shift for a PTM is [0, average\_residue\_length]. Follow the methods described in the spectral probability paper.

(b) The table probabilities in peak\_begin and peak\_end are summed up to obtain precursor probabilities.

\*\*(c) Conditional probabilities are computed based on the ratio (probability above threshold / shift probability \* peak\_width \* peak\_width). The reason for multiplying peak\_width \* peak\_width is the error tolerance.\*\*

\*\*(d) Conditional probabilities is normalized by \* K (0.55) \* 0.25 --> K(0.55) \*\* (remove 0.25)

(e) Count number is the number of modified proteoforms.

(f) adjust factor is 1.0

(g) e-value is conditional\_probability \* count\_number \* adjust\_factor, p-value is 1 - (1-conditional\_prob)^(count\_number \* adjust\_factor)

### 10 Prefix protein-spectrum-matches with two PTMs.

(e) Count number is the number of modified prefixes \* \*\*PREFIX\_SUFFIX\_ADJUST (0.693)\*\*

### 11 Suffix protein-spectrum-matches with two PTMs.

(e) Count number is the number of modified suffixes \* \*\*PREFIX\_SUFFIX\_ADJUST (0.693)\*\*

### 12 Internal protein-spectrum-matches with two PTMs.

(e) Count number is the number of modified internals \* \*\*INTERNAL\_ADJUST (0.508) --> (0.05)\*\*