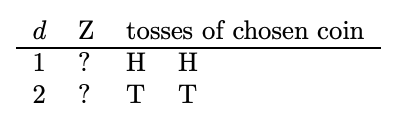
**CSU44062 – Computational Linguistics**

**Assignment 2**

**Brandon Dooley - #16327446**

**Expectation Maximisation (EM)**



As detailed in the assignment we are concerned with the above scenario where a coin Z is tossed to choose between one of the other two coins A and B. The chosen coin is then tossed N (N=2) times and the outcome is recorded in regards to heads (H) or tails (T). In our given scenario the outcome of coin toss Z is hidden and is therefore the *hidden variable*.

Throughout this working we will use the following notation:

num of heads in trial d

num of tails in trial d

a given trial d

To carry out an EM estimation of the parameters given the data we will need some initial setting of the parameters. For this assignment we will suppose this is:

**Iteration 1**

For each piece of data we have to first compute the conditional probabilities of the hidden variable given the data

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Then from these ‘expected’ counts we re-estimate our parameters as follows:

**Iteration 2**

Using our new estimations for our parameters we repeat again for a second iteration with the following values:

For each piece of data we have to first compute the conditional probabilities of the hidden variable given the data

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Then from these ‘expected’ counts we re-estimate our parameters as follows:

Thus at the end of the second iteration our new estimates for the parameters are: