Iterative Bayesian Inference Quality Control

Compute Worker Quality Score

The quality score is the probability prior based on the worker's work. It is sometimes called accuracy, as it correlates strongly with accuracy measures.

$$Q_u = \frac{\sum_{i=0}^{|t|} 1^{A_{u,i}=C_i \wedge D_i>2 \wedge E_i>0.7 \wedge C_i \neq 0 \wedge A_{u,i} \neq 0} + 0.05}{\sum_{i=0}^{|t|} 1^{D_i>2 \wedge E_i>0.7 \wedge C_i \neq 0 \wedge A_{u,i} \neq 0} + 0.5}$$

Determine Task Accuracy

This computation uses the worker quality scores and the agreement/disagreement to compute a final confidence and most-likely value for each task.

- G_{i,g} = Guess Matrix. G_{i,g} represents the probability that task i has answer g
- A_{u,i} = Answer Matrix. A_{u,i} represents the answer for worker u and task i
- Q_u = Quality of worker u
- |W| = number of workers

$$G_{i,g} = \prod_{u=1}^{|W|} 1 - 1^{A_{u,i}=g} * Q_u$$

The answer of highest likelihood for task i is C_i.

The computation of the answer of highest likelihood is given by the following equation.

$$Y_i = \operatorname{argmax}_g \prod_{u=1}^{|W|} 1 + 1^{A_{u,i} \neq g} * 1^{Au,i \neq 0} * (Q_u - 1))$$