(m_1, m_2, \ldots, m_n) in such a way that the procurement cost (that is, cost of annotation), given by \( \sum_{i=1}^{n} m_i c(\eta_i) \), is minimized subject to the PAC learning constraint (3). This is an integer linear programming problem. An approximate solution can be obtained by relaxing the integer constraint and rounding off the optimal solution to the nearest integer value.

Now, let us consider a more realistic scenario of incomplete information where the learner does not know noise rates \( \eta = (\eta_1, \ldots, \eta_n) \). There are two possible approaches: (1) estimation, and (2) elicitation. In the estimation approach, the learner estimates \( \eta_i \) using previously acquired examples (say, for example, comparing labels from different annotators). In the elicitation approach, the learner gets \( \eta_i \) directly from the annotators. The former approach has the disadvantage that poor estimates result in either paying more (when overestimated) or not satisfying the PAC bound (when underestimated). Due to this reason, we are interested in elicitation. In this approach, the learner pays an incentive (a.k.a. price of information) to get \( \eta_i \) from the annotators. Note that the learner needs to pay this price of information to elicit true noise rates. (Otherwise, the annotators can falsely report the noise rate.) For this purpose, we propose to design a procurement auction mechanism to procure a feasible annotation plan with minimum cost; now, the procurement cost also includes the price of information. This problem is challenging because from annotator’s perspective, he would like to maximize his utility (i.e., the payment received minus the internal cost for annotation). The choice of mechanism depends crucially on various design parameters such as \( N, \epsilon, \delta, \eta_i, c(\cdot) \), and the choice of the learning algorithm. We assume that \( N, \epsilon, \delta, c(\cdot) \), and the choice of the learning algorithm are public knowledge, and only \( \eta_i \) is the private information of \( i^{th} \) annotator.

### 3.1 Procurement Auction Model

The learner solicits simultaneous and confidential bids for the noise rates from annotators. Let \( \hat{\eta}_i \) be the bid of \( i^{th} \) annotator that can possibly be a false noise rate. Assume that annotator \( i \) draws his true noise rate \( \eta_i \) in an independent random manner using a density function \( \phi_i \) in the interval \( I_i = [0, 1/3] \) with the corresponding cumulative distribution function \( \Phi_i \), and let \( \phi_i(\eta_i) > 0 \) for all \( \eta_i \in I_i \) and \( i = 1, 2, \ldots, n \). Let \( I = I_1 \times I_2 \times \ldots \times I_n \) and \( \phi = \phi_1 \times \phi_2 \times \ldots \times \phi_n \) denote respective joint spaces. We use the subscript \( -i \) to exclude \( i^{th} \) annotator in any variable (e.g. \( I_{-i}, \eta_{-i} \)) and, we also use the notation \( \hat{\eta} = (\hat{\eta}_1, \ldots, \hat{\eta}_n) \).

After receiving the bids (i.e., \( \hat{\eta} = (\hat{\eta}_1, \ldots, \hat{\eta}_n) \)), the learner allocates a contract of supplying certain number of labeled examples to each annotator and an associated payment. Thus, a procurement auction mechanism is a pair of mappings \( \mathcal{M} = (a, p) \), where \( a : I \mapsto \mathbb{N}_0^n \) is the allocation rule and \( p : I \mapsto \mathbb{R}^n \) is the payment rule.

Given an auction mechanism \( \mathcal{M} = (a, p) \), an annotator \( i \), having noise rate \( \eta_i \), gets the following utility when all the annotators report their bids \( \hat{\eta} \):

\[
 u_i(\hat{\eta}; \eta_i) = p_i(\hat{\eta}) - a_i(\hat{\eta})c(\eta_i) \tag{7}
\]

Note that the first and the second term denote the payment received from the learner and the internal cost in supplying the labeled examples, respectively. Since each annotator \( i \) does not know \( \eta_{-i} \) and moreover, others’ bids \( \hat{\eta}_{-i} \) affect his utility, it is useful to define expected allocation rule \( \alpha \) and the expected payment rule \( \pi \) for any mechanism \( \mathcal{M} = (a, p) \) in the following manner (from \( i^{th} \) annotator’s perspective).

\[
 \alpha_i(\hat{\eta}_i) = \int_{I_{-i}} a_i(\hat{\eta}_i, \hat{\eta}_{-i}) \phi_{-i}(\hat{\eta}_{-i}) d\hat{\eta}_{-i} \tag{8}
\]

\[
 \pi_i(\hat{\eta}_i) = \int_{I_{-i}} p_i(\hat{\eta}_i, \hat{\eta}_{-i}) \phi_{-i}(\hat{\eta}_{-i}) d\hat{\eta}_{-i} \tag{9}
\]

The expected utility of annotator \( i \), when he bids \( \hat{\eta}_i \) while having true value \( \eta_i \), can now be given by

\[
 U_i(\hat{\eta}_i; \eta_i) = \pi_i(\hat{\eta}_i) - \alpha_i(\hat{\eta}_i)c(\eta_i) \tag{10}
\]

When both arguments in (10) are same, we use \( U_i(\eta_i) \) to mean \( U_i(\hat{\eta}_i; \eta_i) \) (for notational simplicity). Given this background, we first present several definitions that are essential to prove our results. A Mechanism \( \mathcal{M} = (a, p) \) is said to be:

\[\text{The symbol } \mathbb{N}_0 \text{ denotes the set of natural numbers inclusive of zero. The allocation and the payment rules are functions of } N, \epsilon, \delta, c(\cdot), \text{ and the algorithm. For notational simplicity, we drop these parameters.}\]