# OOD-ENS: Cross-Validated Out-of-Distribution Ensemble Detector

## A The relationship of $p(p_z)$ and $p_z$.

Before providing the relationship of $p(p_z)$ and $p_z$, we introduce Eq. (7) in the submitted paper as follows.

$$p(p_z) = \int_{0.5}^{1} Be((m_e - 1)p_z, (m_e - 1)(1 - p_z))dx.$$

where $m_e = \frac{K}{1+(K-1)\rho}$ and $\rho = Corr(\hat{g}_k, \hat{g}_{k'})$ with $k, k' = 1, \ldots, K$ and $k \neq k'$.

According to the above formula, we calculate the values of $p(p_z)$ for different $p_z$, $K$, and $\rho$ under the condition of $0 < \rho < 1$ and $3 \leq K \leq n' - 1$, and plot the variance of $p(p_z)$ along with an increases $K$. The results are shown in Figure 1.

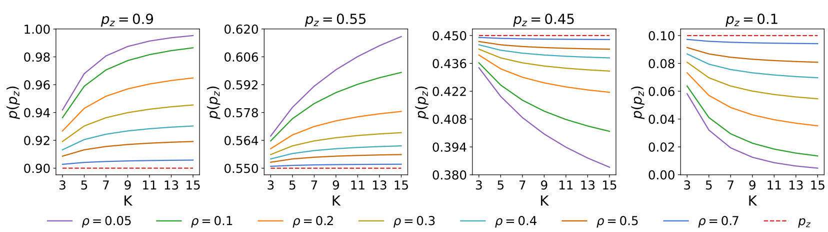


Fig. 1: The relationship of $p(p_z)$ and $p_z$.

From the Figure, it is clear that $p(p_z) > p_z$, and $p(p_z)$ is an increasing function with regard to $K$ when $p_z > 0.5$. On contrary, when $p_z < 0.5$, $p(p_z) < p_z$, and $p(p_z)$ is an decreasing function with regard to $K$. For example, $p(p_z)$ is larger than $0.9$ in the left-top subfigure when $p_z = 0.9$, and $p(p_z)$ is smaller than $0.45$ in the left-bottom subfigure when $p_z = 0.45$. Therefore, the relationship of $p(p_z)$ and $p_z$ is as follows.

– When $p_z > 0.5$, $p(p_z) > p_z$ and $p(p_z)$ is increasing with an increasing $K$.
– When $p_z < 0.5$, $p(p_z) < p_z$ and $p(p_z)$ is decreasing with an increasing $K$.

## B The proof of Lemma 1.

**Lemma 3.1** *For the confusion matrix $\mathcal{H}_\gamma$ of an OOD-ENS algorithm, when $3 \leq K \leq n' - 1$ and $0 < \rho < 1$, the expectations of $TP_\gamma$ and $TN_\gamma$ monotonically increase with an increasing $K$; the expectations of $FP_\gamma$ and $FN_\gamma$ monotonically decrease.*

*Proof.* For the confusion matrix $\mathcal{H}_\gamma$ of an OOD-ENS algorithm, the expectations of $\text{TP}_\gamma$, $\text{FP}_\gamma$, $\text{FN}_\gamma$ and $\text{TN}_\gamma$ are defined as follows.

$$
\begin{aligned}
\mathbb{E}(\text{TP}_\gamma) &= \mathbb{E}\Big(\sum_{j=1}^{n'} \mathbf{1}\big(p(\hat{g}_{oe}=1) \geq \gamma, g_j = 1\big)\Big) \\
&= \sum_{j=1}^{n'} P\big(p(\hat{g}_{oe}=1) \geq \gamma, g_j = 1\big) \\
&= \sum_{j=1}^{n'} P(g_j = 1) \cdot P\big(p(\hat{g}_{oe}=1) \geq \gamma | g_j = 1\big)
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}(\text{FP}_\gamma) &= \sum_{j=1}^{n'} P(g_j = 0) \cdot P\big(p(\hat{g}_{oe}=1) \geq \gamma | g_j = 0\big) \\
\mathbb{E}(\text{FN}_\gamma) &= \sum_{j=1}^{n'} P(g_j = 1) \cdot P\big(p(\hat{g}_{oe}=0) \geq \gamma | g_j = 1\big) \\
\mathbb{E}(\text{TN}_\gamma) &= \sum_{j=1}^{n'} P(g_j = 0) \cdot P\big(p(\hat{g}_{oe}=0) \geq \gamma | g_j = 0\big),
\end{aligned}
$$

where $\hat{g}_{oe}$ denotes the majority-voting prediction and $g_j$ denotes a gold label. Without loss of generality, we assume that an algorithm is at least a weak learning algorithm, which indicates $p_z > 0.5$ for the records of $g_j = 1$, and $p_z < 0.5$ for the records of $g_j = 0$, where $p_z = P(\hat{g}_k = 1)$.

According to the Appendix A, when $3 \leq K \leq n' - 1$ and $0 < \rho < 1$, the following analyses can be obtained.

– For $\text{TP}_\gamma$ and $\text{TN}_\gamma$, we have $P(p(\hat{g}_{oe} = 1) \geq \gamma | g_j = 1)$ and $P(p(\hat{g}_{oe} = 0) \geq \gamma | g_j = 0)$ is an increasing function with regard to $K$.
– For $\text{FP}_\gamma$ and $\text{FN}_\gamma$, we have $P(p(\hat{g}_{oe} = 1) \geq \gamma | g_j = 0)$ and $P(p(\hat{g}_{oe} = 0) \geq \gamma | g_j = 1)$ is a decreasing function with regard to $K$.

On this basis, the confusion matrix $\mathcal{H}_\gamma$ has the optimal properties that the expectations of $\text{TP}_\gamma$ and $\text{TN}_\gamma$ monotonically increase with an increasing $K$; the expectations of $\text{FP}_\gamma$ and $\text{FN}_\gamma$ monotonically decrease.

## C The proof of Theorem 1

**Theorem 3.2** *Under the condition of* $0 < \rho < 1$*, the expectation of the evaluation measure ACC of an OOD-ENS algorithm increases with an increasing* $K$ *and owns an upper bound when* $K$ *equals to* $n' - 1$*, while those of estimators* $\alpha$ *and* $\beta$ *decrease and own an lower bound.*

*Proof.* The expectations of the evaluation measure ACC, the Type I error $\alpha$ and the type II error $\beta$ of an OOD-ENS algorithm are defined as follows.

$$
\begin{aligned}
\mathbb{E}(\text{ACC}) &= \mathbb{E}\left(\frac{\text{TP}_\gamma + \text{TN}_\gamma}{n'}\right) = \frac{\mathbb{E}(\text{TP}_\gamma) + \mathbb{E}(\text{TN}_\gamma)}{n'},\\
\mathbb{E}(\alpha) &= \mathbb{E}\left(\frac{\text{FP}_\gamma}{\text{FP}_\gamma + \text{TN}_\gamma}\right) = \frac{\mathbb{E}(\text{FP}_\gamma)}{\sum_j \mathbf{1}(g_j = 0)},\\
\mathbb{E}(\beta) &= \mathbb{E}\left(\frac{\text{FN}_\gamma}{\text{FN}_\gamma + \text{TP}_\gamma}\right) = \frac{\mathbb{E}(\text{FN}_\gamma)}{\sum_j \mathbf{1}(g_j = 1)},
\end{aligned}
$$

where $n' = \text{TP}_\gamma + \text{FP}_\gamma + \text{FN}_\gamma + \text{TN}_\gamma$ is a constant and $\mathbf{1}(\cdot)$ is an indicator function.

Under the condition of $0 < \rho < 1$, according to the Lemma 1, the expectations of ACC increases with an increasing $K$ and owns an upper bound when $K$ equals to $n' - 1$, while those of estimators $\alpha$ and $\beta$ decrease and own an lower bound.