Yingzhen Li (yingzhen.li@imperial.ac.uk)

# Short tutorial on risk-based definition of uncertainty

## 0. Preliminary: Commonly used uncertainty measures

Suppose we are in supervised learning setting, and we have an underlying data distribution $p_d(y|x)$ generating the dataset of $N$ datapoints $\mathcal{D} = \{(x_i, y_i)\}_{i\in[N]}$. The machine learning approach starts from making a model $p(y|x,\theta)$ with some parameter $\theta \in \Theta$ to be fitted using the dataset $\mathcal{D}$. Here we are interested in not only the prediction accuracy, but also uncertainty estimates from the machine learning model.

### 0.1. Popular uncertainty measures

Specifically, assuming the ground truth is not deterministic, this means $p_d(y|x)$ is not a delta distribution, and therefore we can consider uncertainty measures to evaluate e.g., how "spread-out" the data distribution is.

Examples of such uncertainty measures include (the two most popular ones):

- **Variance:** $\mathbb{V}_{p_d(y|x)}[y] := \mathbb{E}_{p_d(y|x)}[(y - \mathbb{E}_{p_d(y|x)}[y])^2]$.
- **Entropy:** $\mathbb{H}[p_d(y|x)] := -\mathbb{E}_{p_d(y|x)}[\log p_d(y|x)]$.

### 0.2. Model specification and Bayesian estimation

In practice we do not know $p_d(y|x)$ but rather only has an estimate of it. By using model estimates, we need to consider two possible sources of error:

- **Model misspecification:** is there a "true" $\theta^* \in \Theta$ such that $p_d(y|x) = p(y|x,\theta^*)$?
- **Finite data estimation error:** even with no model specification, point estimate $\hat{\theta}$ coming from the finite dataset $\mathcal{D}$ is unlikely to satisfy $p(y|x,\hat{\theta}) = p(y|x,\theta^*)$.

For now we will assume there is no model error and focus on the estimation error.

With no model error, we can assume $p_d(y|x) = p(y|x, \theta^*)$ for a "ground truth" $\theta^* \in \Theta$, which means the "true aleatoric uncertainty" (or true data uncertainty) can be defined using $p(y|x, \theta^*)$.

Examples of such true aleatoric uncertainty measures include (the two most popular one):

- Variance: $\mathbb{V}_{p_d(y|x)}[y] = \mathbb{V}_{p(y|x,\theta^*)}[y]$.
- Entropy: $\mathbb{H}[p_d(y|x)] = \mathbb{H}[p(y|x, \theta^*)]$.

However, we do not know $\theta^*$ but instead we only have a finite dataset $\mathcal{D} = \{(x_i, y_i)\}_{i \in [N]}$ with $y_i | x_i \sim p(y|x, \theta^*)$. There are two options to estimate uncertainty then:

- Option 1: fit point estimate, e.g., $\hat{\theta} = \hat{\theta}_{\text{MLE}}$, and then compute e.g., variance as $\mathbb{V}_{p_d(y|x)}[y] \approx \mathbb{V}_{p(y|x,\hat{\theta})}[y]$.
- Option 2: fit a distribution estimate $q(\theta)$ based on data and then consider the following estimates (with variance as an example):
    - $\mathbb{V}_{p_d(y|x)}[y] \approx \mathbb{E}_{q(\theta)}[\mathbb{V}_{p(y|x,\theta)}[y]]$
    - $\mathbb{V}_{p_d(y|x)}[y] \approx \mathbb{V}_{q(y|x)}[y], q(y|x) := \int p(y|x, \theta) q(\theta) d\theta$

Which estimate should we choose? In fact we can consider both estimates, if $q(\theta)$ is a "reasonable" distribution satisfying contraction properties, i.e., $q(\theta) \to \delta(\theta = \theta^*)$ as $N \to +\infty$ with good rates (e.g., variance of $q(\theta)$ contracts in a rate of $\mathcal{O}(N^{-1})$).

- Bayesian posterior $p(\theta|\mathcal{D})$ is a good choice when the prior $p(\theta)$ is chosen such that $\theta^* \in supp(p(\theta))$ (refer to e.g., Bayesian CLT results).

Now let's revisit the two potential choices with $q(\theta) = p(\theta|\mathcal{D})$ (under the variance example):

- $\mathbb{V}_{p_d(y|x)}[y] \approx \mathbb{E}_{p(\theta|\mathcal{D})}[\mathbb{V}_{p(y|x,\theta)}[y]]$: this can be viewed as a Bayesian estimate of the aleatoric uncertainty (variance). From Bayesian CLT, as

$p(\theta|\mathcal{D}) \to \delta(\theta = \theta^*)$ when $N \to +\infty$, we have the true aleatoric uncertainty recovered: $\mathbb{E}_{p(\theta|\mathcal{D})}[\mathbb{V}_{p(y|x,\theta)}[y]] \to \mathbb{V}_{p_d(y|x)}[y]$.

- $\mathbb{V}_{p_d(y|x)}[y] \approx \mathbb{V}_{p(y|x,\mathcal{D})}[y], p(y|x,\mathcal{D}) := \int p(y|x,\theta)p(\theta|\mathcal{D})d\theta$: this can be viewed as thr "total uncertainty" estimate, where the prediction is done by averaing all possible fits $p(y|x,\theta), \theta \sim p(\theta|\mathcal{D})$ to produce the prediction. From Bayesian CLT, this estimate will also converge to the true aleatoric uncertainty.

The interesting thing is that the "total uncertainty" (TU) is bigger than the "Bayesian estimate of aleatoric uncertainty" (AU). It is easy to show for e.g., variance:

$$\mathbb{V}_{p(y|x,\mathcal{D})}[y] = \mathbb{E}_{p(\theta|\mathcal{D})}[\mathbb{V}_{p(y|x,\theta)}[y]] + \mathbb{V}_{p(\theta|\mathcal{D})}[\mathbb{E}_{p(y|x,\theta)}[y]].$$

This decomposition suggest the additive uncertainty decomposition idea which defines the concept of "epistemic uncertainty" (EU):

$$\mathrm{TU} = \mathrm{AU} + \mathrm{EU}.$$

In the. case of variance as uncertainty measure, the EU definition becomes

$$\mathrm{EU} = \mathbb{V}_{p(\theta|\mathcal{D})}[\mathbb{E}_{p(y|x,\theta)}[y]].$$

Similarly, for entropic measure of uncertainty, we can also construct such decomposition:

- $\mathrm{TU} = \mathbb{H}[p(y|x,\mathcal{D})]$, where $p(y|x,\mathcal{D}) := \int p(\theta|\mathcal{D})p(y|x,\theta)d\theta$.
- $\mathrm{AU} = \mathbb{E}_{p(\theta|\mathcal{D})}[\mathbb{H}[p(y|x,\theta)]]$.
- $\mathrm{EU} = \mathbb{E}_{p(\theta|\mathcal{D})}[\mathrm{KL}[p(y|x,\theta)||p(y|x,\mathcal{D})]]$.

In fact, this decomposition applies beyond variance and entropic definition of uncertainty measures, and generalised to risk-based uncertainty measures as described below.

# 1. Minimal Excess Risk as a Generalised Epistemic Uncertainty Measure

We discuss a generalised definition of Epistemic Uncertainty (EU) via Bayesian decision theory. Most of the below content contains known results but just rewritten in a different way. See https://arxiv.org/abs/2012.14868 for an example reference for further reading.

## 1.1. Decision-making via (Bayes) risk minimisation

We consider possible input states $x \in X$, outputs $y \in Y$ and actions $a \in A$. Assume that nature's law of predicting $y$ given $s$ is captured by the following model with parameter $\theta^*$:

$$p(y|x, \theta^*). \quad \text{(ground-truth predictive law)}$$

Assuming each possible input-output-action triplet $(x, y, a)$ is associated with a loss function $l(a, y, x)$. Then in a hypothetical world, assuming $\theta^*$ is known, we can compute the expected loss ("risk") of an action $a \in A$ as:

$$R_l(a, x, \theta^*) := \mathbb{E}_{p(y|x,\theta^*)}[l(a, y, x)].$$

Therefore, the best decision which minimises the expected loss under the nature's law is:

$$a^*(x, \theta^*) := \arg\min_{a \in A} \mathbb{E}_{p(y|x,\theta^*)}[l(a, y, x)] = \arg\min_{a \in A} R_l(a, x, \theta^*).$$

This means under the optimal decision according to nature's law, there is also an irreducible amount of loss, which is also named irreducible risk:

$$R_l(x, \theta^*) := \min_{a \in A} \mathbb{E}_{p(y|x,\theta^*)}[l(a, y, x)] = \mathbb{E}_{p(y|x,\theta^*)}[l(a^*(x, \theta^*), y, x)].$$

In practice, however, we do not know $\theta^*$, but we may have some "belief" $q(\theta)$ about the value of $\theta$. For example, with a dataset of $N$ datapoints $\mathcal{D} = \{(x_i, y_i)\}_{i \in [N]}$, a Bayesian approach would define a prior $p(\theta)$, and then obtain a posterior $p(\theta|\mathcal{D})$ using the likelihood model $p(y|x, \theta)$ and use $p(\theta|\mathcal{D})$ as the "belief" $q(\theta)$. In general, with a "belief" $q(\theta)$, we can compute an ``belief estimate'' of the irreducible risk as:

$$R_l(x, q) := \mathbb{E}_{q(\theta)}[R(x, \theta)] = \mathbb{E}_{q(\theta)}[\min_{a \in A} \mathbb{E}_{p(y|x,\theta)}[l(a, y, x)]].$$

However, the "belief estimate" of irreducible risk does not suggest a single action to perform, which is less ideal. To address this issue, we can also estimate the risk of each individual action $a \in A$ as:

$$R_l(a, x, q) := \mathbb{E}_{q(\theta)}[R(a, x, \theta)] = \mathbb{E}_{q(\theta)p(y|x,\theta)}[l(a, y, x)].$$

This also suggest a "belief-optimal" action which minimises the above risk estimate:

$$a^*(x, q) := \arg\min_{a \in A} \mathbb{E}_{q(\theta)p(y|x,\theta)}[l(a, y, x)] = \arg\min_{a \in A} R_l(a, x, q).$$

Again the optimal "expected loss" here is non-zero, which we name as belief-optimal decision risk:

$$\tilde{R}_l(x, q) := \min_{a \in A} \mathbb{E}_{q(\theta)p(y|x,\theta)}[l(a, y, x)] = \mathbb{E}_{p(\theta)p(y|x,\theta)}[l(a^*(x, q), y, x)].$$

In below notes, when setting $q(\theta) := p(\theta|\mathcal{D})$, we also use the following notations:

$$R_l(x, \mathcal{D}) := R_l(x, p(\theta|\mathcal{D})), \quad R_l(a, x, \mathcal{D}) := R_l(a, x, p(\theta|\mathcal{D})),$$
$$a^*(x, \mathcal{D}) := a^*(x, p(\theta|\mathcal{D})), \quad \tilde{R}_l(x, \mathcal{D}) := \tilde{R}_l(x, p(\theta|\mathcal{D})).$$

In such case, we also change the terminology of "belief-optimal" to "Bayes-optimal". To relate to the known concepts in Bayesian decision theory, $R_l(a, x, \mathcal{D})$ is the Bayes risk for a given action $a$, and $\tilde{R}_l(x, \mathcal{D})$ represents the minimum Bayes risk. In fact $R_l(a, x, q)$ is also generally referred as the Bayes risk for a given action $a$ in the literature, however here we explicitly state this term as "q-belief" to emphasise the fact that $q$ may not correspond to a Bayesian posterior belief.

## 1.2. Minimum Excess Risk (MER) as a generalised epistemic uncertainty for a given belief

An important proposition which will be used later is the following:

**Prop 1.1 (Generalised data processing inequality).** $\tilde{R}_l(x,q) \geq R_l(x,q)$, with equality achived iff. $q(\theta)$ is a delta measure, or $a^*(x,\theta) := \arg\min_{a\in A} \mathbb{E}_{p(y|x,\theta)}[l(a,y,x)]$ is constant w.r.t. $\theta$.

**Proof.** using Jensen's inequality applied to $\min(\cdot)$ function:

$$\begin{aligned} \tilde{R}_l(x,q) &:= \min_{a\in A} \mathbb{E}_{q(\theta)}[\mathbb{E}_{p(y|x,\theta)}[l(a,y,x)]] \\ &\geq \mathbb{E}_{q(\theta)}[\min_{a\in A} \mathbb{E}_{p(y|x,\theta)}[l(a,y,x)]] =: R_l(x,q). \end{aligned}$$

This also means we can define the Minimum Excess Risk (MER) of optimal decision under belief $q(\theta)$ as:

$$\mathrm{MER}_l(x,q) := \tilde{\mathcal{R}}_l(x,q) - R_l(x,q) \geq 0.$$

We can interpret the defined three terms above as a generalised version of total, aleatoric and epistemic uncertainties (TU, AU and EU) for a given belief $q(\theta)$, quantified using a loss function $l(a,y,x)$:

- $\mathrm{TU}_l(x,q) := \tilde{R}_l(x,q)$.
    - The "belief-optimal" decision $a^*(x,q)$ takes the $q$ belief in all possible "nature laws" $\theta \sim q(\theta)$ into account when making a decision. This single decision $a^*(x,q)$ is not optimal for any possible $\theta \sim q(\theta)$, but it has on average the lowest risk when averaged over all possible $\theta \sim q(\theta)$.
- $\mathrm{AU}_l(x,q) := R_l(x,q)$.
    - As aleatoric uncertainty quantifies the inherent randomness in nature which is irreducible, $R_l(x,q)$ as a "belief estimate" of irreducible risk can also be viewed as an "belief estimate" of true aleatoric uncertainty quantified using a loss function.
- $\mathrm{EU}_l(x,q) := \mathrm{MER}_l(x,q) := \tilde{R}_l(x,q) - R_l(x,q)$.
    - Assuming optimal decision $a^*(x,\theta)$ is indeed dependent on $\theta$, we see that $\mathrm{EU}_l(x,q) = 0$ iff. $q(\theta)$ reaches to a delta measure (thus no belief uncertainty on $\theta$ anymore). Since the beliefs are often regarded as "subjective" (e.g., posterior belief $p(\theta|\mathcal{D})$), this term can also be regarded as an epistemic uncertainty measure.

These definitions satisfy the additive decomposition rule: $\mathrm{TU} = \mathrm{AU} + \mathrm{EU}$.

## 1.3. Examples

We now present some examples on how the above definitions of TU, AU and EU recovers the widely used versions based on e.g., entropy and variance. We also show some additional examples related to some interesting applications.

**Example 1. Variance**

We define $A = Y$ (i.e., action and output spaces coincide) and use the L2 loss as the loss function $l(a, y, x) := (y - a)^2$:

- $\min_{a \in A} \mathbb{E}_{p(y|x,\theta)}[l(a, y, x)] = \min_{a \in A} \mathbb{E}_{p(y|x,\theta)}[(y - a)^2] = \mathbb{V}_{p(y|x,\theta)}[y]$.
- $\mathrm{AU}_l(x, q) = \mathbb{E}_{q(\theta)}[\mathbb{V}_{p(y|x,\theta)}[y]]$.
- $\mathrm{EU}_l(x, q) = \mathbb{V}_{q(\theta)}[\mathbb{E}_{p(y|x,\theta)}[y]]$.
- $\mathrm{TU}_l(x, q) = \mathbb{V}_{q(y|x)}[y]$, where $q(y|x) := \int q(\theta) p(y|x, \theta) d\theta$.

This decomposition agrees with the law of total variance.

When using $q(\theta) := p(\theta|\mathcal{D})$, we also have:

- $\mathrm{AU}_l(x, \mathcal{D}) = \mathbb{E}_{p(\theta|\mathcal{D})}[\mathbb{V}_{p(y|x,\theta)}[y]]$.
- $\mathrm{EU}_l(x, \mathcal{D}) = \mathbb{V}_{p(\theta|\mathcal{D})}[\mathbb{E}_{p(y|x,\theta)}[y]]$.
- $\mathrm{TU}_l(x, \mathcal{D}) = \mathbb{V}_{p(y|x,\mathcal{D})}[y]$.

**Example 2. Stochastic action, and Entropy & mutual information**

To account for the inherent stochasticity in $p(y|x, \theta)$, one can also make a “stochastic action” via a policy $\pi(a)$, in this case we can also define a loss $l(\pi, y, x)$ on policy level, and then the “risk” definition can also be extended to a policy $\pi$:

$$R_l(\pi, x, \theta) := \mathbb{E}_{p(y|x,\theta)}[l(\pi, y, x)].$$

This means we can also minimise the risk over a policy and extend the minimum risk definitions to stochastic actions draw from policy $\pi \in \Pi$. With an abuse of notation (reloading), we write:

$$\pi^*(x, \theta) := \arg\min_{\pi \in \Pi} \mathbb{E}_{p(y|x,\theta)}[l(\pi, y, x)],$$

$$R_l(x, \theta) := \min_{\pi \in \Pi} \mathbb{E}_{p(y|x,\theta^*)}[l(\pi, y, x)],$$

$$R_l(x, q) := \mathbb{E}_{q(\theta)}[R_l(x, \theta)] = \mathbb{E}_{q(\theta)}[\min_{\pi \in \Pi} \mathbb{E}_{p(y|x,\theta^*)}[l(\pi, y, x)]],$$

$$\tilde{R}_l(x, q) := \min_{\pi \in \Pi} \mathbb{E}_{q(\theta)p(y|x,\theta^*)}[l(\pi, y, x)].$$

Now we show that Prop 0.1, stated under optimal deterministic action case, can be extended to stochastic action case:

**Prop 1.2.** $\tilde{R}_l(x, q) \geq R_l(x, q)$, with equality achived iff. $q(\theta)$ is a delta measure, or $\pi^*(x, \theta) := \arg\min_{\pi \in \Pi} \mathbb{E}_{p(y|x,\theta)}[l(\pi, y, x)]$ is constant w.r.t. $\theta$.

The proof is essentially the same as the proof of Prop. 0.1.

This also means we can carry on with the definition of MER, TU, AU, EU as in the deterministic action space.

Now consider again $A = Y$ and define the policy-level loss as $l(\pi, y, x) := -\log \pi(y)$. Then we notice that the “risk” becomes the cross-entropy loss:

$$\begin{aligned} R_l(\pi, x, \theta) &:= \mathbb{E}_{p(y|x,\theta)}[-\log \pi(y)] \\ &:= \mathrm{CE}(p(y|x,\theta), \pi(y)) = \mathrm{KL}[p(y|x,\theta)||\pi(y)] + \mathbb{H}[p(y|x,\theta)]. \end{aligned}$$

This also returns the following uncertainty definitions, assuming $\Pi$ contains all possible distributions over $y \in Y$:

- $\min_{\pi \in \Pi} \mathbb{E}_{p(y|x,\theta)}[l(\pi, y, x)] = \mathbb{H}[p(y|x,\theta)]$.
- $\mathrm{AU}_l(x, q) = \mathbb{E}_{q(\theta)}[\mathbb{H}[p(y|x,\theta)]]$.
- $\mathrm{EU}_l(x, q) = \mathbb{E}_{q(\theta)}[\mathrm{KL}[p(y|x,\theta)||q(y|x)]]$, where $q(y|x) := \int q(\theta)p(y|x,\theta)d\theta$.
- $\mathrm{TU}_l(x, q) = \mathbb{H}[q(y|x)]$, where $q(y|x) := \int q(\theta)p(y|x,\theta)d\theta$.

When using $q(\theta) := p(\theta|\mathcal{D})$, we also have:

- $\mathrm{AU}_l(x, \mathcal{D}) = \mathbb{E}_{p(\theta|\mathcal{D})}[\mathbb{H}[p(y|x,\theta)]]$.
- $\mathrm{EU}_l(x, \mathcal{D}) = \mathbb{E}_{p(\theta|\mathcal{D})}[\mathrm{KL}[p(y|x,\theta)||q(y|x)]] = \mathbb{I}[y; \theta|\mathcal{D}]$.
- $\mathrm{TU}_l(x, \mathcal{D}) = \mathbb{H}[p(y|x,\mathcal{D})]$.

This decomposition agrees with the relationship between total entropy, conditional entropy and mutual information.

### Example 3. Semantic Uncertainty

The semantic uncertainty paper considered entropic definitions of uncertainty measures but with an additional step: grouping the outputs with similar semantic meaning. This can also be framed into the generalised EU definition via a particular loss definition.

In detail, let's assume that the $Y$ space can be split into disjoint semantic classes $Y = \cup_{c \in C} Y_c$ with $Y_c \neq \emptyset, \forall c \in C$ and $Y_c \cap Y_{c'} = \emptyset, c \neq c' \in C$. Now consider the action space as $A = C$ as well as stochastic actions (i.e., treat $c$ as random and define policy $\pi(c) \in \Pi$). We now use the following loss function in risk calculations:

$$l(\pi, y, x) = -\log \pi(c)\delta(y \in Y_c).$$

Then notice:

$$\begin{aligned} R_l(\pi, x, \theta) &:= \mathbb{E}_{p(y|x,\theta)}[-\log \pi(c)\delta(y \in Y_c)] = \mathbb{E}_{p(c|x,\theta)}[-\log \pi(c)] \\ &:= \mathrm{CE}(p(c|x,\theta), \pi(c)) = \mathrm{KL}[p(c|x,\theta)||\pi(c)] + \mathbb{H}[p(c|x,\theta)], \end{aligned}$$

with $p(c|x,\theta) := \int p(y|x,\theta)\delta(y \in Y_c)dy$.

This also returns the following uncertainty definitions, assuming $\Pi$ contains all possible distributions over $c \in C$:

- $\min_{\pi \in \Pi} \mathbb{E}_{p(y|x,\theta)}[l(\pi, y, x)] = \mathbb{H}[p(c|x,\theta)]$.
- $\mathrm{AU}_l(x, q) = \mathbb{E}_{q(\theta)}[\mathbb{H}[p(c|x,\theta)]]$.
- $\mathrm{EU}_l(x, q) = \mathbb{E}_{q(\theta)}[\mathrm{KL}[p(c|x,\theta)||q(c|x)]]$, where $q(c|x) := \int q(\theta)p(c|x,\theta)d\theta$.
- $\mathrm{TU}_l(x, q) = \mathbb{H}[q(c|x)]$, where $q(c|x) := \int q(\theta)p(c|x,\theta)d\theta$.

When using $q(\theta) := p(\theta|\mathcal{D})$, we also have:

- $\mathrm{AU}_l(x, \mathcal{D}) = \mathbb{E}_{p(\theta|\mathcal{D})}[\mathbb{H}[p(c|x,\theta)]]$.
- $\mathrm{EU}_l(x, \mathcal{D}) = \mathbb{E}_{p(\theta|\mathcal{D})}[\mathrm{KL}[p(c|x,\theta)||q(y|x)]] = \mathbb{I}[c; \theta|\mathcal{D}]$.
- $\mathrm{TU}_l(x, \mathcal{D}) = \mathbb{H}[p(c|x,\mathcal{D})]$.

This decomposition agrees with the relationship between total entropy, conditional entropy and mutual information.

**Example 4. Bayesian quantile regression and value-at-risk (VaR)**

Quantile regression considers $A = Y$ and a pinball loss for a given a quantile level $\tau \in [0, 1]$:

$$\rho_\tau(a, y, x) := (y - a)(\tau - \delta(y \leq a)).$$

This returns the following uncertainty estimators:

- $\min_{a \in A} \mathbb{E}_{p(y|x,\theta)}[\rho_\tau(a, y, x)] = \mathrm{CDF}^{-1}_{p(y|x,\theta)}(\tau) := \inf\{y : P(Y \leq y|x, \theta) \geq \tau\}$.
    - This is the $\tau$-quantile of the distribution $p(y|x, \theta)$.
- $\mathrm{AU}_{\rho_\tau}(x, q) = \mathbb{E}_{q(\theta)}[\mathrm{CDF}^{-1}_{p(y|x,\theta)}(\tau)]$.
- $\mathrm{TU}_{\rho_\tau}(x, q) = \mathrm{CDF}^{-1}_{q(y|x)}(\tau)$, where $q(y|x) := \int q(\theta)p(y|x, \theta)d\theta$.
- $\mathrm{EU}_{\rho_\tau}(x, q) = \mathrm{CDF}^{-1}_{q(y|x)}(\tau) - \mathbb{E}_{q(\theta)}[\mathrm{CDF}^{-1}_{p(y|x,\theta)}(\tau)]$.

In particular, if the $Y$ variable represents some loss values (e.g., in finance applications), $\mathrm{AU}_{\rho_r}(x, q)$ can be viewed as a belief-estimate of the value-at-risk (VaR) at quantile level $\tau$.

**Example 5: Per-action loss estimation with uncertainty**

Interestingly, we can also define the following “meta-level” minimum risk to estimate the loss of a particular action with uncertainty.

Specifically, we define $l_a := l(a, y, x)$ for $y \sim p(y|x, \theta)$ as the loss function $l$ evaluated on a given action $a$. Since $y$ is a random variable, it means that $l_a$ is also a random variable induced by $y$, and we can write down its distribution as $p(l_a|x, \theta)$.

We provide 2 examples below to illustrate the idea of per-action risk estimation with uncertainty.

Quantile estimation:

Now consider estimating the quantile of the loss function $l_a$ evaluated on a given action $a$. Using the minimum risk framework at meta level, we can define a “meta-action” $\tilde{a} \in \mathbb{R}$ (since $l_a \in \mathbb{R}$) and consider the pinball loss as the “meta loss” for a “meta minimum risk” defined as follows:

$$R_{\rho_\tau}(x, a, \theta) := \min_{\tilde{a} \in \mathbb{R}} \mathbb{E}_{p(l_a|x,\theta)}[\rho_\tau(\tilde{a}, l_a, x)] = \min_{\tilde{a} \in \mathbb{R}} \mathbb{E}_{p(y|x,\theta)}[\rho_\tau(\tilde{a}, l(a, y, x), x)].$$

We note here that the definition is also on a given $a$, since $l_a$ can be viewed as a conditional random variable conditioned on both $x$ and $a$. In other words, the input space at this meta level is $(x, a) \in X \times A$.

This returns the following uncertainty estimators:

- $R_{\rho_\tau}(x, a, \theta) = \mathrm{CDF}^{-1}_{p(l_a|x,\theta)}(\tau) := \inf\{l : P(l(a, Y, x) \leq l | x, \theta) \geq \tau\}$.
    - This is the $\tau$-quantile of the distribution $p(l_a|x, \theta)$.
- $\mathrm{AU}_{\rho_\tau}(x, a, q) = \mathbb{E}_{q(\theta)}[\mathrm{CDF}^{-1}_{p(l_a|x,\theta)}(\tau)]$.
- $\mathrm{TU}_{\rho_\tau}(x, a, q) = \mathrm{CDF}^{-1}_{q(l_a|x)}(\tau)$, where $q(l_a|x) := \int q(\theta) p(l_a|x, \theta) d\theta$.
- $\mathrm{EU}_{\rho_\tau}(x, a, q) = \mathrm{CDF}^{-1}_{q(l_a|x)}(\tau) - \mathbb{E}_{q(\theta)}[\mathrm{CDF}^{-1}_{p(l_a|x,\theta)}(\tau)]$.

Variance estimation:

Similarly, we can also use the L2 loss at meta level $l_2(\tilde{a}, l_a, x) := (l_a - \tilde{a})^2$:

- $R_{l_2}(x, a, \theta) = \min_{\tilde{a} \in \mathbb{R}} \mathbb{E}_{p(y|x,\theta)}[(l(a, y, x) - \tilde{a})^2] = \mathbb{V}_{p(y|x,\theta)}[l(a, y, x)]$.
- $\mathrm{AU}_{l_2}(x, a, q) = \mathbb{E}_{q(\theta)}[\mathbb{V}_{p(y|x,\theta)}[l(a, y, x)]]$.
- $\mathrm{EU}_{l_2}(x, a, q) = \mathbb{V}_{q(\theta)}[\mathbb{E}_{p(y|x,\theta)}[l(a, y, x)]]$.
- $\mathrm{TU}_{l_2}(x, a, q) = \mathbb{V}_{q(y|x)}[l(a, y, x)]$, where $q(y|x) := \int q(\theta) p(y|x, \theta) d\theta$.

This means if we define a reward function as $r(a, x) := -l(y, a, x), y \sim p(y|x, \theta)$ (which is again a random variable, with the ground truth reward sampled using $p(y|x, \theta^*)$), the classical UCB algorithm can be written as (per step):

$$a^*_{UCB}(x, q) := \arg\min_{a \in A} \mathbb{E}_{q(y|x)}[l(a, y, x)] - \beta \sqrt{\mathrm{EU}_{l_2}(x, a, q)}.$$