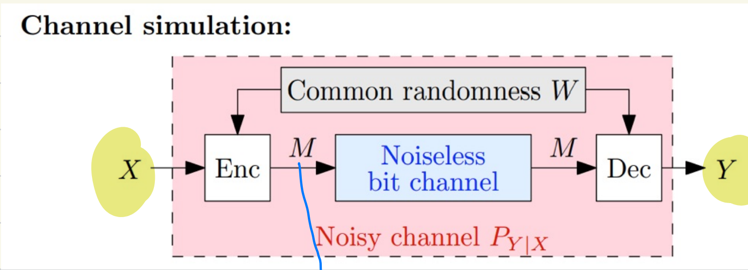
3.4. Likelihood Encoder

Recap : Channel simulation



$\sim P_{Y|X}$

$L := \mathbb{E}[|M|]$. Question: $L^* = \inf L$.

Gumbel-Max trick

Goal: generate samples following a given probability distribution.

Given: $a_1, a_2, \dots, a_n > 0$.

Generate $z_i \sim \exp(r_i)$. Let $\frac{z_i}{a_i/r_i} := \phi(z_i)$

$$P\Big(\phi(z_y) \le \underset{j \neq y}{\phi(z_j)}\Big) = \frac{a_y}{\sum a_i}$$

when $\sum a_i = 1 \Rightarrow P\left(\arg\min \phi(z_i) = y\right) = a_y$.

time index of samples

Likelihood sampling: allow estimation of mean of a function of a distribution P when we can only access sample from another distribution Q.

Suppose we have i.i.d. samples $\overline{Y}_1, \overline{Y}_2, \dots, \overline{Y}_N$ following the reference distribution Q. We want to estimate $\mathbb{E}_{Y\sim P}[f(Y)]$ for a target distr. P.

Then:

$$E_{Y\sim P}[f(Y)] \approx \sum_{i=1}^{N} f(\overline{Y}_i) \frac{\frac{dP}{dQ}(\overline{Y}_i)}{\sum_{i=1}^{N} \frac{dP}{dQ}(\overline{Y}_i)}$$

Notes:

1. $\frac{dP}{dQ}$: Radon-Nikodym derivative.

When P, Q discrete $\rightarrow \dfrac{P(\overline{Y}_i)}{Q(\overline{Y}_i)}$

2. $\approx$ is approximation. " $=$ " in the asymptotic regime $N \rightarrow \infty$.

3. Let $\dfrac{dP}{dQ}(\overline{Y}_i) := \alpha_i$ (importance weight)

$$E_{Y \sim P}[f(Y)] \approx \sum_{i=1}^{N} f(\overline{Y}_i) \frac{\frac{dP}{dQ}(Y_i)}{\sum_{i=1}^{N} \frac{dP}{dQ}(\overline{Y}_i)}$$

$$= \sum_{i=1}^{N} f(\overline{Y}_i) \frac{\alpha_i}{\sum \alpha_i}$$

Gumbel-Max trick.

$\rightarrow$ Minimal random coding scheme

1. Reference distribution Q and $\overline{Y}_1, \overline{Y}_2, \ldots, \overline{Y}_N \overset{iid}{\sim} Q$ known to both encoder and decoder.

2. Encoder observes input X, wants to convey sample $\bar{Y} \sim P = P_{Y|X}(\cdot|X)$ to decoder.

① Compute importance weights $\alpha_i := \frac{dP}{dQ}(\bar{Y}_i)$

② Generate numbers $K \in [N]$ with $P(K=k) = \frac{\alpha_k}{\sum \alpha_i}$

③ Send the binary representation of K ($\approx \log_2 N$ bits)

If source distribution P_X is known, take $Q = P_Y$.

If X, Y discrete, then:

$$\alpha_i = \frac{dP}{dQ}(\bar{Y}_i) = \frac{P(\bar{Y}_i)}{Q(\bar{Y}_i)} = \frac{P_{Y|X}(\bar{Y}_i|X)}{P_Y(\bar{Y}_i)}$$

$$\underset{\text{Bayes's rule}}{=} \frac{\overbrace{P_{X|\bar{Y}}(X|\bar{Y}_i)}^{\text{likelihood}}}{\underbrace{P_X(X)}_{\text{known}}}$$

Approximation guarantee:

$N \to \infty$: exact $P_{Y|X}$ simulation

$N \approx 2^{D_{KL}(P\|Q)}$: 'good' approximation

Theorem [Liu, Verdú | 2018]. For any sampling scheme on $\bar{Y}_1, \bar{Y}_2, \ldots, \bar{Y}_N \overset{iid}{\sim} Q$ with output distribution $\bar{Y}_K \sim P$, we have:

$$\boxed{E[K] \geq 2^{D_{KL}(P\|Q)-1}}$$

3.5. Comparison between sampling schemes.

1. Rejection sampling

- Exactly $Y \sim P$ simulation.
- $L^* \leq I(X;Y) + \log_2(I(X;Y) + \log_2(4e)) + \log_2(8e)$
- Requires computations of expectations, which might be impossible.

2. Poisson Functional representation:

- Exactly $Y \sim P$
- $L^* \leq I(X; Y) + \log_2 (I(X; Y) + 2) + 3$

3. Minimal random coding and Likelihood encoder.

- Approximate: $Y \sim P$ as $N \to \infty$.

- May or may not achieve $L^* = I(X; Y)$.

Achieve in asymptotic regime $N \to \infty$