

The code length for arithmetic coding

- ▶ Given m bits of binary expansion we assume the rest are all zero.
- ▶ Distance between two m bit expansions is 2^{-m}
- ▶ If $l_T - u_T \geq 2^{-m}$ then there must be a point x described by m expansion bits such that $l_T \leq x < u_T$
- ▶ Required number of bits is $\lceil -\log_2(u_T - l_T) \rceil$.
- ▶ $u_T - l_T = \prod_{t=1}^T p(c_t | c_1, c_2, \dots, c_{t-1}) \doteq p(c_1, \dots, c_T)$
- ▶ Number of bits required to code c_1, c_2, \dots, c_T is $\lceil -\sum_{t=1}^T \log_2 p_t(c_t) \rceil$.
- ▶ We call $-\sum_{t=1}^T \log_2 p_t(c_t) = -\log_2 p(c_1, \dots, c_T)$ the Cumulative log loss