Sampling the Categorical Distribution

$$W \sim Cat(\underline{\theta}) = \prod_{d=0}^{D-1} \theta_d^{I(d=w)}$$

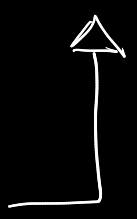
How to sample?

↳ How to create a dataset

$W \in \{Cloudy^{0}, Rainy^{1}, Sunny^{2}\}$

given $\underline{\theta}$

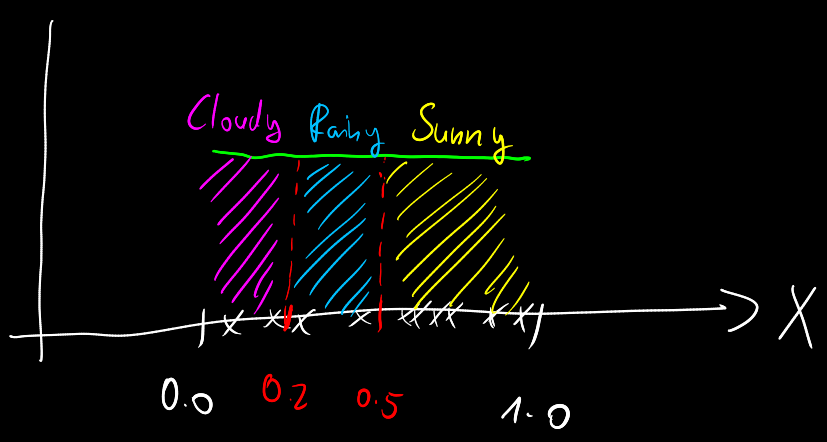
get $D = \{0, 0, 2, 1, 0, 1, 1\}$



Sampling

e.g. : $\underline{\theta} = \begin{bmatrix} 0.2 \\ 0.3 \\ 0.5 \end{bmatrix}$ → Cloudy, → Rainy, → Sunny

Source of Randomness : Uniform Distribution



$X \sim Uniform(0.0, 1.0)$

Uniform → Categorical

$\underline{q} = \begin{bmatrix} 0.2 \\ 0.5 \\ 1.0 \end{bmatrix}$ $\quad q_i = \sum_{j=0}^{i} \theta_j$ → cummulated sum

↳ cumsum

Example:

$D_X = \{0.4, 0.9, 0.3, 0.1, 0.7, 0.5, 0.2, \dots\}$

$D_W = \{1, 2, 1, 0, 2, 1, 0, \dots\}$