

A vector of
model parameters

A vector of
sufficient statistics

$$\Pr(\mathbf{Y} = \mathbf{y} \mid \theta, \mathbf{X}) = \frac{\exp \{ \theta^t s(\mathbf{y}, \mathbf{X}) \}}{\sum_{\mathbf{y}' \in \mathcal{Y}} \exp \{ \theta^t s(\mathbf{y}', \mathbf{X}) \}}, \quad \forall \mathbf{y} \in \mathcal{Y}$$

Observed data

The normalizing
constant

All possible
networks