

A convergence criterion would be when there is no more significant increase in the lower bound, as defined by $\text{align}^* = \int q^*(z, \lambda, \alpha) \log [p(z, \lambda, \alpha) q^*(z, \lambda, \alpha)] dz$.

Note that the categorical pmf $p(y_i | y_{ij}^*)$ becomes degenerate once the latent variables are known, so this term is cancelled out. With the exception of $q^*(z)$, all of the distributions are Gaussian. The following results will be helpful.

definition[Differential entropy] The differential entropy of a pdf $p(x)$ is given by $H(p) = -\int p(x) \log p(x) dx = -\mathbb{E}_p[\log p(x)]$.